



Universidad
Politécnica
de Cartagena

TRABAJO FIN DE MÁSTER

GRADUADO EN INGENIERÍA DE SISTEMAS DE
TELECOMUNICACIÓN

Aplicación de modelos de aprendizaje profundo a segmentación de imágenes IVUS

Carlos García Navarro



Director: Juan Morales Sánchez

Departamento de Tecnologías de la Información y las Comunicaciones

14 de septiembre de 2023

Agradecimientos

En primer lugar, gracias a Ángela por ser un apoyo constante e incondicional.

Gracias a mi tutor, Juan Morales, por su paciencia durante la realización de este Trabajo de Fin de Máster.

Gracias a Shengnan y Gijs por hacer posible mi estancia en Rotterdam. Gracias a mis compañeros de Erasmus MC: Antonio, Nuria, Jonas, Aaron, Sowmiya y Anton.

También doy las gracias a mis compañeros de Universidad por los buenos momentos compartidos.

Resumen

Este trabajo aborda la segmentación simultánea de los contornos del lumen y el vaso sanguíneo en imágenes de IVUS, lo que podría suponer una valiosa ayuda para los cardiólogos en la toma de decisiones durante la Intervención Coronaria Percutánea. Con este propósito, se escogió un modelo de *Deep Learning*, Deeplabv3+, con *Modified Alligned Xception* como columna vertebral de la red, y se utilizaron datos del ensayo *FFR react trial*. Además, se aplicaron técnicas avanzadas tales como aumento de datos, optimización de hiperparámetros mediante búsqueda bayesiana y postprocesado de las predicciones mediante morfología matemática para mejorar las prestaciones del flujo de trabajo elegido. La implementación final optimizada obtuvo, para el caso de imágenes sin *stent*, una puntuación Dice de 0,952 y 0.960 para el lumen y la pared vascular, respectivamente, y una distancia Hausdorff de 0.207 mm y 0.278 mm para el contorno del lumen y la pared vascular, respectivamente. En el caso de imágenes con *stent*, el lumen obtuvo un coeficiente de Dice de 0.967 y una Distancia Hausdorff de 0.242 mm, mientras que la pared vascular obtuvo unos valores de 0.940 y 0.455 mm, respectivamente. Esto demuestra que el enfoque aquí presentado es capaz de realizar una detección robusta y precisa del contorno del lumen y la pared vascular en imágenes IVUS.

Palabras clave: Deep Learning; segmentación; IVUS; lumen; pared vascular; *stent*; *pull-back*.

Abstract

This work addresses the simultaneous segmentation of lumen and vessel contours in IVUS images, which could be a valuable assist for cardiologists in decision making during Percutaneous Coronary Intervention. For this purpose, a Deep Learning model, Deeplabv3+, with Modified Alligned Xception as backbone of the network was chosen, and data from the FFR react trial were used. Additionally, advanced techniques such as data augmentation, Bayesian search hyperparameter optimization and post-processing using mathematical morphology were applied to improve the performance of the chosen framework. The final optimized implementation obtained, for the case of images with no stent, a Dice score of 0.952 and 0.960 for the lumen and vascular wall, respectively, and a Hausdorff distance of 0.207 mm and 0.278 mm for the lumen and vascular wall contour, respectively. In the case of images with stent, the lumen obtained a Dice coefficient of 0.967 and a Hausdorff Distance of 0.242 mm, while the vascular wall obtained values of 0.940 and 0.455 mm, respectively. This proves that the approach presented here is capable of robust and accurate detection of the lumen and vascular wall contours in IVUS images.

Keywords: Deep Learning; segmentation; IVUS; lumen; vessel; stent; pullback.

Índice general

1. Introducción	1
1.1. Enfermedad coronaria del corazón	1
1.2. Modalidades de imagen intravascular	2
1.2.1. Estructura del vaso sanguíneo	2
1.2.2. IVUS	3
1.2.3. OCT	5
1.2.4. NIRS	5
1.3. <i>State-of-the-art</i>	7
1.4. Objetivos	9
2. Métodos	10
2.1. Modelo de Deep Learning	10
2.1.1. Deeplabv3+	10
2.1.2. Modified Alligned Xception	11
2.1.3. Función de pérdidas, optimizador y activación de la última capa	12
2.2. <i>Bayesian Search</i>	13
2.2.1. Función Objetivo	13
2.2.2. Función subrogada	13
2.2.3. Función de adquisición	14
2.3. <i>Data augmentation</i>	15
2.3.1. Deformación elástica	16
2.3.2. Deformación polar	17
2.4. Morfología matemática	18
2.4.1. Erosión	19
2.4.2. Dilatación	19
2.4.3. Apertura	20
2.4.4. Cierre	20
3. Materiales y métricas	22
3.1. Base de Datos	22

3.2. Funciones de I/O	22
3.2.1. Manejo de <i>data.Dataset</i>	23
3.3. Metodología	24
3.4. Métricas	24
4. Resultados	28
4.1. Vecindario	28
4.2. Selección de los hiperparámetros con <i>Bayesian Search</i>	30
4.3. Postprocesado	32
4.4. <i>Stent</i>	35
4.5. Comparativa con otros métodos	38
5. Discusión	41
5.1. Futuras líneas	43
Referencias	49

Índice de figuras

1.1.	PCI desde la angiografía.	2
1.2.	Esquemático del corte transversal de la arteria.	3
1.3.	Imagen ejemplo de <i>pullback</i> y cortes transversales de IVUS en pre-intervención.	4
1.4.	Imagen ejemplo de <i>pullback</i> y cortes transversales de IVUS en post-intervención.	4
1.5.	Comparativa entre imágenes de OCT (A y C) y de IVUS (B y D) para una misma sección transversal de arteria.	6
1.6.	Ejemplo de imagen NIRS.	7
1.7.	Ejemplo de NIRS-IVUS. Las zonas amarillas se corresponden con placa con lípidos y la roja con zonas donde no existe placa con lípidos.	8
2.1.	Estructura de Deeplabv3+.	11
2.2.	Esquemático de la convolución dilatada.	11
2.3.	Esquema de <i>Modified Aligned Xception</i> . Los cambios realizados por los autores se encuentran en naranja.	12
2.4.	Ejemplo función de adquisición.	14
2.5.	Proceso global de <i>Bayesian Search</i>	14
2.6.	Evolución del proceso de optimización en función de la adquisición seleccionada.	15
2.7.	Comportamiento de deformación elástica en función de α	16
2.8.	Funcionamiento de la deformación polar.	18
2.9.	Ejemplos de distintos elementos estructurantes.	18
2.10.	Ejemplo de erosión morfológica.	19
2.11.	Ejemplo de dilatación morfológica.	19
2.12.	Ejemplo de apertura morfológica.	20
2.13.	Ejemplo de cierre morfológico	21
3.1.	Flujograma del manejo de los datos.	23
3.2.	Conclusiones extraíbles de los valores de AUC.	25
3.3.	Intersección y unión de conjuntos.	26
3.4.	Ejemplo de medida de Distancia Hausdorff	26

4.1.	Contornos de la predicción al utilizar una única imagen e imágenes multiespectrales de 3 canales como <i>input</i> del modelo.	29
4.2.	Imágenes con sobresegmentación y contornos irregulares sin postprocesar.	32
4.3.	Diagramas de violín para el <i>fold</i> 3 de las distintas métricas.	34
4.4.	Contornos de predicciones sin <i>stent</i> del modelo final optimizado.	35
4.5.	Diagramas de violín para el <i>fold</i> 2 para las distintas métricas en imágenes con <i>stent</i>	37
4.6.	Contornos de predicciones con <i>stent</i> del modelo final optimizado.	38

Índice de tablas

4.1. Métricas de los cuatro enfoques seleccionados.	28
4.2. Métricas antes y después de la optimización bayesiana.	30
4.3. Comparación entre las deformaciones elástica y polar.	31
4.4. Métricas de tres etiquetas distintas con datos sin <i>stent</i>	33
4.5. Métricas de las tres etiquetas en imágenes con <i>stent</i>	36
4.6. Comparativa entre distintos métodos.	39

Glosario

Adam	Adaptative Moment Stimation
ASPP	Atrous Spatial Pyramid Pooling
AUC	Area Under the Curve
DA_{th}	Data Augmentation threshold
EI	Expected Improvement
DSC	Sørensen–Dice Coefficient
FFR	Fractional Flow Reserve
FoV	Field of View
GPR	Gaussian Process Regresor
GPU	Graphics Processor Unit
HD	Hausdorff Distance
IoU	Intersection over Union
IVUS	Intravascular Ultrasound
LCB	Lower Confidence Bound
LSTM	Long Short-Term Memory
NIRS	Near Infrared Spectroscopy
OCT	Optical Coherence Tomography
OS	Output Stride
PCI	Percutaneous Coronary Intervention
PI	Probability of Improvement
RAM	Random Access Memory
ReLU	Rectified Linear Unit
ROI	Region Of Interest
ROC	Receiver Operating characteristic Curve
UCB	Upper Confidence Bound

1 Introducción

1.1. Enfermedad coronaria del corazón

La enfermedad coronaria del corazón es el tipo de enfermedad cardíaca más común. Es un estrechamiento o bloqueo de las arterias coronarias, que resulta en un aprovisionamiento insuficiente de flujo sanguíneo al corazón, pudiendo finalizar en un ataque cardíaco [1]. Se estima que 126 millones de personas son diagnosticadas con enfermedad coronaria del corazón en todo el mundo [2], siendo la principal causa de muerte en Estados Unidos [3].

Esta enfermedad se debe a la acumulación de colesterol y otros elementos, que se recogen bajo el nombre de placa, en la capa interna de las paredes de la arteria. Esta acumulación recibe el nombre de arterioesclerosis [3].

El procedimiento principal a la hora de tratar esta enfermedad es la Intervención Percutánea Coronaria (PCI), que consiste en la apertura de la arteria por medio de un globo o el posicionamiento de una malla metálica denominada *stent* en la zona objetivo de la arteria [4]. El *stent* impide a los lípidos crecer en el interior del lumen, de manera que se evite bloquear el flujo de sangre en la arteria, pudiendo, eso sí, desarrollarse alrededor del mismo.

Durante esta intervención, la forma más habitual de proceder es utilizar simultáneamente angiografía y algún tipo de imagen intravascular. Esto permite tanto localizar la zona a la que el riego sanguíneo no llega, como ver cuál es la causa interna de esta falta de riego mediante la imagen intravascular.

Generalmente, la zona objetivo de la arteria es explorada varias veces durante la intervención [4]: se realiza una exploración previa, que sirve para comprobar si es necesario hacer algún tipo de tratamiento; una post-intervención tras una angiografía correcta, en la que se asegura haber situado el *stent* correctamente; y se explora una última vez para comprobar si se puede optimizar de alguna manera el posicionamiento del *stent*. El uso de la angiografía permite ver el flujo global de sangre y comprobar que no sea necesario situar un *stent* debido a que hay otra zona taponada que no se veía anteriormente [5]. En la Figura 1.1

se tiene un ejemplo ilustrativo de este proceso mediante la angiografía.

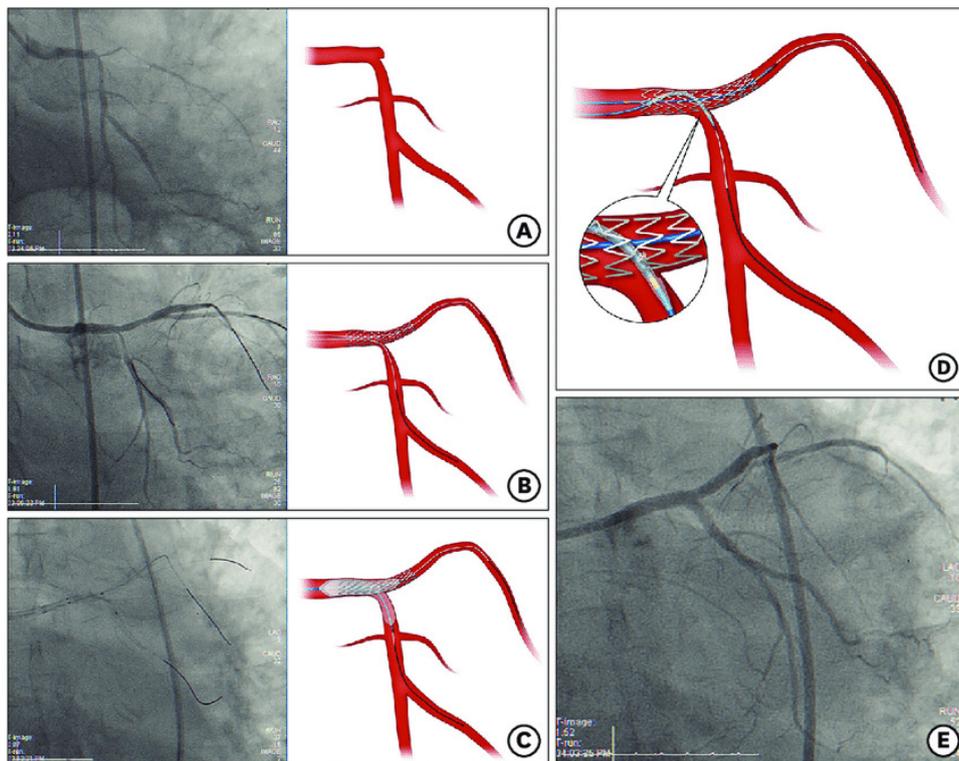


Figura 1.1: PCI desde la angiografía. Fuente: [5].

En adición a la angiografía, durante este proceso es fundamental tener una imagen intravascular que permita visualizar y asegurar que ninguna zona ha sido dañada durante la intervención o que se haya pasado por alto algún cuerpo extraño que pudiera ser problemático en el futuro.

1.2. Modalidades de imagen intravascular

1.2.1. Estructura del vaso sanguíneo

Se considera esencial en este trabajo comprender la estructura transversal de un vaso sanguíneo para conocer qué estructuras pueden estar afectando al comportamiento del algoritmo en la segmentación.

En una arteria se encuentran tres capas: *Intima*, *Media* y *Adventitia*. En la Figura 1.2 se puede ver un esquemático incluyendo dichas capas y otras estructuras que pueden encontrarse en el corte transversal de una sección de la arteria. El lumen, por donde circula el flujo sanguíneo, se encuentra delimitado por la *Intima*, y la pared vascular por el interfaz *Media-Adventitia*. Los cristales de colesterol, placa y calcificaciones se sitúan generalmente en la *Intima*.

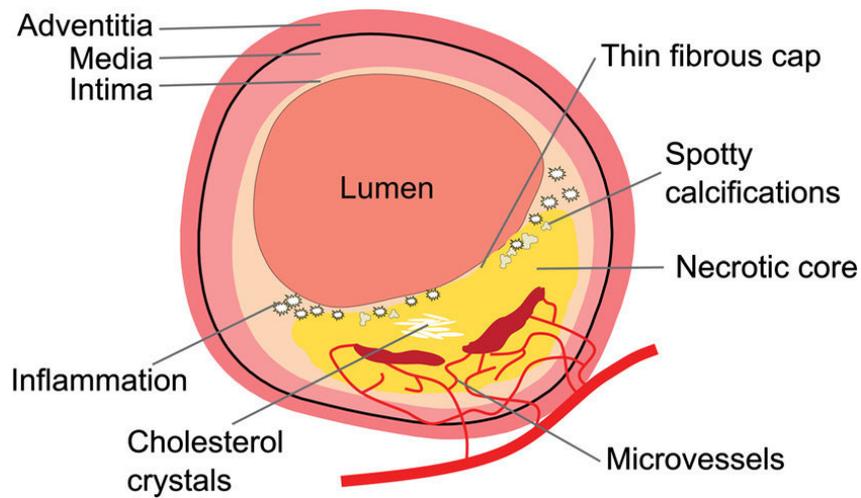


Figura 1.2: Esquemático del corte transversal de la arteria. Fuente: [6].

Aunque no está representado en la Figura 1.2, es posible encontrar trombosis de distinta naturaleza dentro del lumen, siendo este bastante peligroso debido a su dificultad de detección [4].

1.2.2. IVUS

La modalidad de imagen intravascular por excelencia es *Intravascular Ultrasound (IVUS)* [7]. Este es un tipo de imagen adquirido introduciendo un catéter por las arterias y llevándolo hasta una zona objetivo, para allí realizar la adquisición de la imagen. El transceptor, situado en la cabeza del catéter, emite ondas de ultrasonidos, típicamente de 20 MHz o 40 MHz, aunque los más avanzados emiten ondas de 60 MHz. Estas ondas rebotan en los distintos tejidos que componen la arteria y vuelven a ser captados por el transceptor. La imagen se adquiere haciendo rotar el catéter mientras se tira de él, lo que le da nombre a la modalidad de imagen adquirida: *pullback*.

Para la introducción del catéter, generalmente se realiza una incisión en la arteria radial o femoral, por la que se introduce la sonda y se lleva a las arterias coronarias para conocer el nivel de lípidos en dichos vasos sanguíneos. El control y monitoreo de las arterias coronarias es crítico debido a la importancia de estas en el transporte del flujo sanguíneo a través del sistema circulatorio humano, y la rápida evolución que puede tener la placa de lípidos.

Según se trate de pre-intervención o post-intervención, podremos distinguir distintos cuerpos en las imágenes de IVUS. En las Figuras 1.3 y 1.4 se tienen tanto el corte longitudinal de una arteria como distintos cortes transversales de IVUS a lo largo del volumen objetivo en planos con distintas características.

En la Figura 1.3, que se corresponde con un escenario de preintervención, hay distintos elementos visibles en IVUS: en la Figura 1.3a se observa una placa calcificada (indicado por

el asterisco) que se introduce en el lumen (indicado por las flechas). En la Figuras 1.3b y 1.3c se observan dos tipos distintos de trombos. El trombo es el principal enemigo a la hora de segmentar imágenes de IVUS y el hecho de ser detectado previamente ayudaría también a los cardiólogos a la hora de tratar la lesión de una manera concreta. En la Figura 1.3d se observa ruptura de placa (flecha), en la que una parte del lumen (el asterisco) se encuentra rodeada casi completamente por calcio. Por último, en la Figura 1.3e se encuentra placa blanda con atenuación.

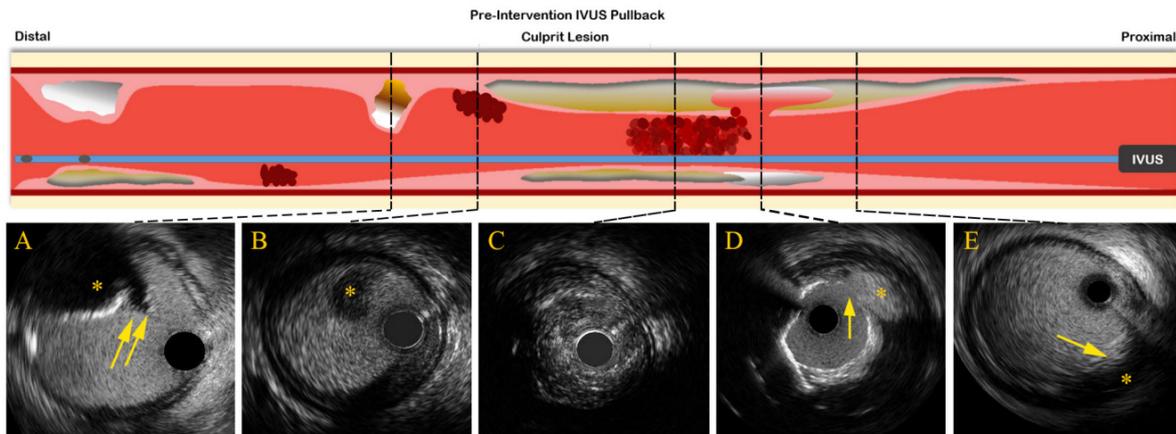


Figura 1.3: Imagen ejemplo de *pullback* y cortes transversales de IVUS en pre-intervención. Fuente: [4].

Por otro lado, las imágenes relacionadas con postintervención se encuentran en la Figura 1.4. Este escenario se sitúa después del posicionamiento del *stent*, que evita que no siga estrechando la arteria por acumulación de lípidos. Las Figuras 1.4a, 1.4b y 1.4c se corresponden con zonas en las que no se ha situado el *stent*. En la Figura 1.4a se ve que las placas calcificadas (flecha) y la fibrosa (asterisco) provocan un estrechamiento del lumen. En la Figura 1.4b se puede ver un hematoma que comprime el lumen (asterisco) y una cierta acumulación de sangre (flecha). En la Figura 1.4e se ve una gran placa lipoadiposa.

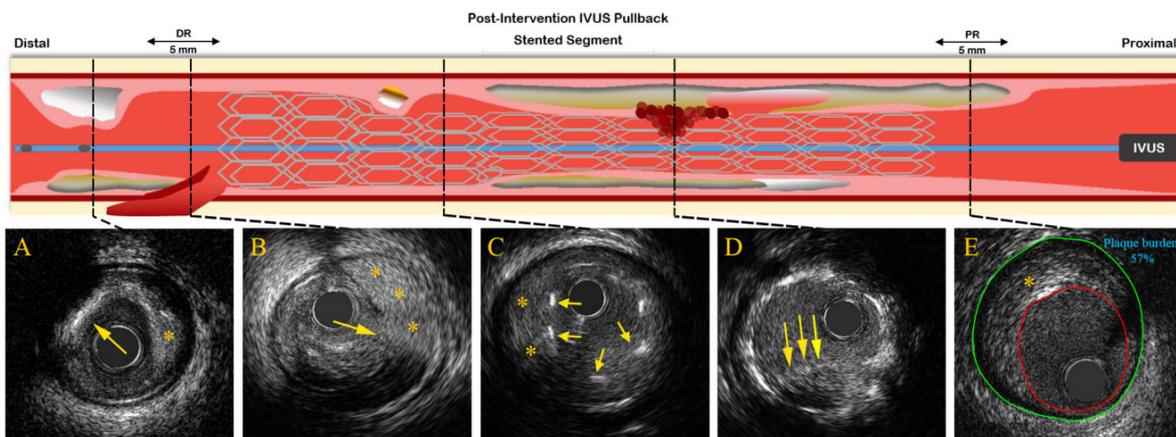


Figura 1.4: Imagen ejemplo de *pullback* y cortes transversales de IVUS en post-intervención. Fuente: [4].

Por otro lado, en las regiones en las que sí se ha situado el *stent* podemos encontrar otro tipo de situaciones. En la Figura 1.4c se ve el mal posicionamiento del *stent*, que provoca que aún haya espacio entre la malla metálica y la pared del lumen, lo que podría provocar el crecimiento de lípidos dentro del lumen a pesar de haber realizado una intervención. En la Figura 1.4d se ve uno de los problemas que se comenta anteriormente; al situar el *stent*, el trombo se queda en el interior del lumen, con la posibilidad de crecer en el interior del mismo. Esto concluiría en la final obstrucción de la circulación sanguínea, y debería ser detectado por el intervencionista cuanto antes.

Los distintos tipos de placa que hemos visto anteriormente son las que provocan el cierre progresivo de la arteria, llevando a un consecuente bloqueo del riego sanguíneo.

1.2.3. OCT

La Tomografía por Coherencia Óptica (OCT) es otra modalidad de imagen utilizada en este ámbito, y que consigue un mejor contraste a la hora de encontrar los límites del lumen y la pared vascular. Esto hace relativamente simple la segmentación de estas regiones en comparación con IVUS. OCT utiliza frecuencias de infrarrojo cercano, en concreto las longitudes de onda de 1250 nm a 1350 nm, y funciona comparando la diferencia entre la señal emitida y recibida en términos de dispersión [8].

OCT es capaz de detectar el mal posicionamiento del *stent*, bifurcaciones y trombosis con gran precisión. Sin embargo, IVUS permite ver con mejor detalle las estructuras que se encuentren en el interior del lumen, lo que hace de este la modalidad de imagen intravascular de facto en PCI [7]. En la Figura 1.5 tenemos un ejemplo de imágenes de IVUS y de OCT, de manera que se pueda diferenciar las características aquí comentadas.

Las fortalezas de IVUS, unidas con las de OCT, y la diferencia de frecuencias utilizada por ambas modalidades, posibilita el uso sincrónico de ambas tecnologías. Diversos trabajos han sido presentados en este sentido, como [10] o [11], y hay gran interés en el desarrollo de esta modalidad.

1.2.4. NIRS

Near Infrared Spectroscopy (NIRS) es una técnica que se basa en la propiedad de las sustancias de absorber o dispersar luz de radiación infrarroja cercana (longitudes de onda de 800 a 2500 nm) a diferencias intensidades en función de la longitud de onda [12].

El uso principal de esta técnica es la detección de placas de lípidos en arterias humanas. A diferencia de las modalidades anteriores, NIRS proporciona una cierta probabilidad de presencia de lípidos y la dirección en la que estos se encuentran, sin poder proporcionar ningún tipo de información sobre la profundidad a la que se encontrarían [12]. En la imagen

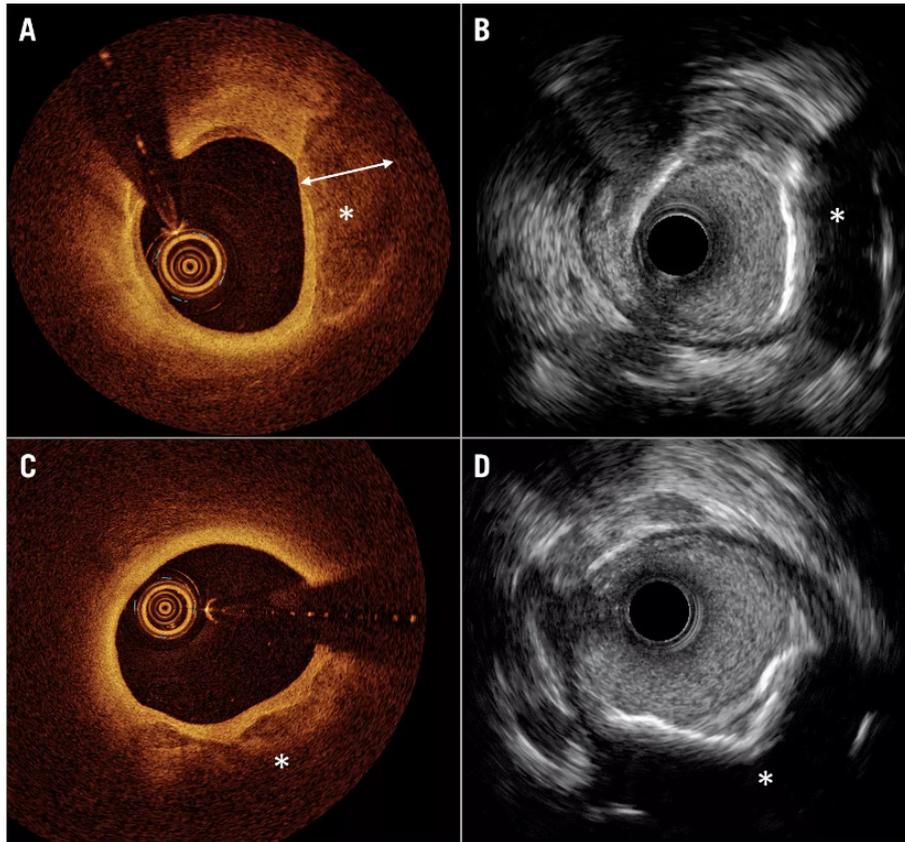


Figura 1.5: Comparativa entre imágenes de OCT (A y C) y de IVUS (B y D) para una misma sección transversal de arteria. Fuente: [9].

1.6 se tiene un ejemplo de la adquisición de NIRS, y su capacidad para detectar placa lípida. Si el NIRS detecta que hay placa lípida en una cierta dirección, aparecerá esta en color amarillo, marcando su presencia. En caso contrario, si no se encuentra placa lípida, se mostrará en color rojo [13].

Esta modalidad de imagen, por sí sola, no es capaz de proporcionar información real sobre el contenido del corte transversal del vaso sanguíneo, lo que la hace realmente limitada [12]. Es por ello que diversos trabajos se han focalizado en unir la información proporcionada por el NIRS con IVUS, como [14], [12] o [15]. Los resultados obtenidos por esta nueva modalidad son muy satisfactorios para conocer si la placa visible en la imagen de IVUS contiene lípidos o no. En la imagen 1.7 se tiene un ejemplo de la unión de estas dos modalidades.

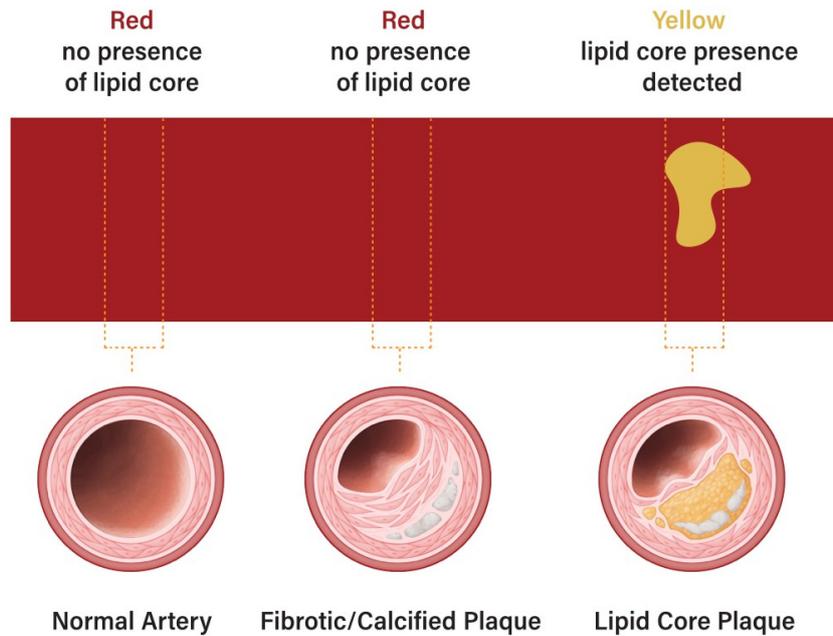


Figura 1.6: Ejemplo de imagen NIRS. Fuente: [13].

1.3. *State-of-the-art*

Los métodos de segmentación tradicionales, tales como umbralización, o métodos más avanzados como los de morfología matemática, en muchos casos no son lo suficientemente robustos para segmentar imágenes médicas debido a su incapacidad de captar características no globales de estas [16]. Además, en la literatura no se encuentran aplicadas a segmentación de imágenes de IVUS, que son en las que atienen a este trabajo. Es por ello que se tratará este problema mediante técnicas de *Machine Learning*, y más concretamente de *Deep Learning*.

En primer lugar, hay trabajos basados en técnicas de *Machine Learning*, como el análisis BLOB [17], que se basa en encontrar distintas regiones de la imagen basándose en brillo o intensidad, para después aplicar análisis de componentes conectados para detectar los límites del lumen y la pared vascular.

Por otra parte, un trabajo centrado en aprendizaje no supervisado [18] presentó unas prestaciones considerables en segmentación de lumen y pared vascular. El algoritmo utilizado en este se deriva del conocido K-means, y se usa una restricción en el tamaño del vecindario para decrementar los requerimientos de cómputo.

En los últimos años, las técnicas de *Deep Learning* son utilizadas principalmente en problemas de segmentación de imágenes médicas [19] debido a sus altas prestaciones y robustez. Se han explorado distintos enfoques, como redes neuronales convolucionales, redes de cápsulas o incluso arquitecturas específicas para segmentación de imágenes médicas o de IVUS.

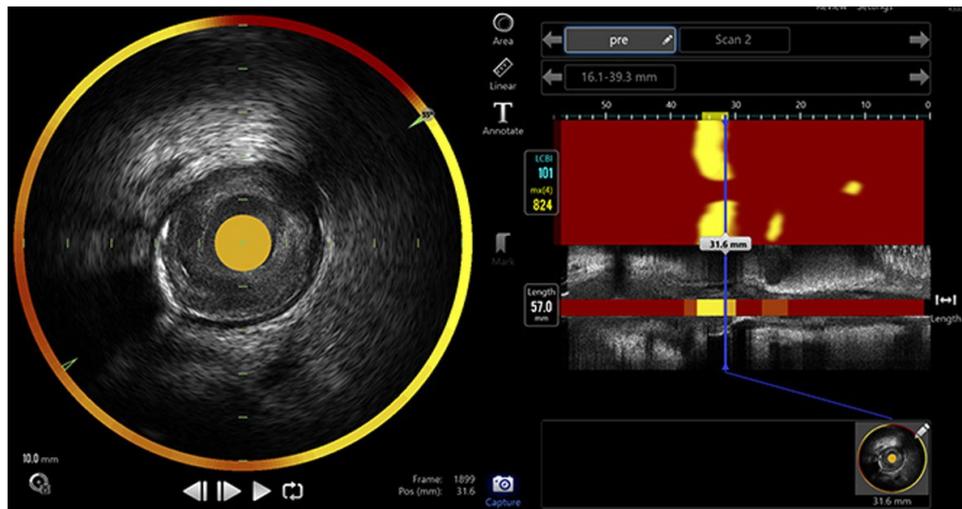


Figura 1.7: Ejemplo de NIRS-IVUS. Las zonas amarillas se corresponden con placa con lípidos y la roja con zonas donde no existe placa con lípidos. Fuente: [14].

En otro trabajo notable [20], los autores aplicaron una red convolucional multicapa y un Regresor de Procesos Gaussianos (*Gaussian Process Regresor* GPR) para segmentar el lumen y eliminar el ruido de alta frecuencia en el contorno de la predicción, respectivamente.

También se encontraron varios trabajos que utilizaron U-Net para segmentar el lumen y la pared vascular. Esta red fue diseñada específicamente para segmentación de imágenes médicas y es considerada un estándar en el campo. En un primer estudio [21], esta red es seguida por un GPR para eliminar el ruido de alta frecuencia en los contornos. De manera adicional, se utilizó la información adyacente de los *frames* del vecindario para interpolar las etiquetas y crear de esta manera más datos para el entrenamiento. Un *frame* se considera a cada una de las divisiones dentro de un *pullback*. En nuestro caso, a cada una de los cortes bidimensionales que se puede encontrar dentro de la imagen tridimensional. En un segundo trabajo [22], se utilizó una U-Net de 8 capas junto con una nueva técnica de aumento de datos para resolver la segmentación de lumen y pared vascular en imágenes de IVUS.

Trabajos basados en redes del estado del arte en materia de segmentación de imágenes naturales también han sido utilizados. Las redes de cápsulas [23], en las que los elementos básicos son neuronas agrupadas en tensores (llamados cápsulas), mostraron unas elevadas prestaciones cuando hay pocas imágenes disponibles para el entrenamiento. IVUS-Net [24], una red diseñada para la segmentación de imágenes de IVUS que incorpora aspectos de distintas redes de segmentación general, presentó unos resultados relevantes en cuanto a segmentación se refiere.

Por último, un trabajo estudió varias de las redes que se comentan anteriormente con una base de datos compacta [25], encontrando que Deeplabv3+ junto con ResNet 101 conseguía mejores resultados que el resto de modelos en cuanto a segmentación de imágenes de IVUS se refiere.

1.4. Objetivos

Teniendo en cuenta todos los conceptos aquí explicados, este trabajo se centrará en la segmentación de imágenes de IVUS mediante algoritmos de aprendizaje profundo, intentando aplicar técnicas que no aplicadas previamente en segmentación de esta modalidad de imagen. También se tomarán ideas de distintos trabajos de manera que se alcance un modelo relevante para ayudar a los cardiólogos durante la PCI.

Se utilizarán distintas métricas para valorar los resultados, de manera que esta comparativa no sea meramente gráfica, y estas métricas servirán para compararlos con los distintos trabajos disponibles en la literatura. Esto permitirá comprobar la bondad de los mismos dentro del contexto del problema, aunque sea sólo de manera ilustrativa, teniendo en cuenta que los datos para entrenar los modelos son distintos.

2 Métodos

2.1. Modelo de Deep Learning

Para el objetivo principal de este proyecto, es decir, la segmentación de lumen y pared vascular en imágenes de IVUS, se propuso el uso de un modelo de aprendizaje profundo previamente utilizado por numerosos trabajos en segmentación de imágenes médicas [19]; Deeplabv3+ con Modified Alligned Xception como columna vertebral del modelo [26].

2.1.1. Deeplabv3+

Deeplabv3+ es un modelo de segmentación de imágenes con propósito general. Para utilizar adecuadamente esta última versión, Deeplabv3+ debe ser situado al final de un modelo de clasificación de imágenes [26]. A este modelo de clasificación se le ha quitado la última capa, de manera que se obtenga un mapa de características como salida de este modelo en lugar de un único valor numérico. Así, Deeplabv3+ es capaz de convertir la información extraída por el modelo de clasificación en una segmentación.

La estructura total de la red se puede ver en la Figura 2.1. La red convolucional sería donde se sitúa Modified Alligned Xception en nuestro caso. Deeplabv3+ se fundamenta en 3 pilares principales para segmentar adecuadamente [26]:

- *Atrous convolution* (o convolución dilatada): es una variación de la convolución 2D en la que se sitúan ceros en las posiciones no nulas del filtro, de manera que se pueda controlar el campo de visión (*Field of View, FoV*) del algoritmo. En la Figura 2.2 se encuentra una descripción gráfica del funcionamiento de la convolución dilatada:

$$y[i] = \sum_{k=1}^K x[i + rk] \cdot \omega[k] \quad (1)$$

- *Atrous Spatial Pyramid Pooling* (ASPP): utiliza la convolución dilatada de manera que identifique objetos a distintas escalas modificando el parámetro r de la ecuación ante-

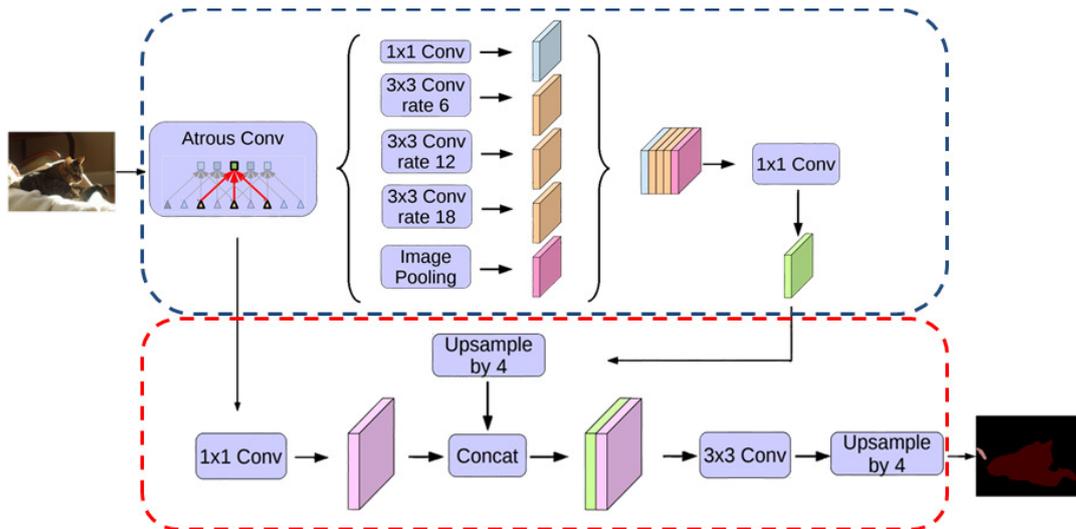


Figura 2.1: Estructura de Deeplabv3+. Fuente: [26].

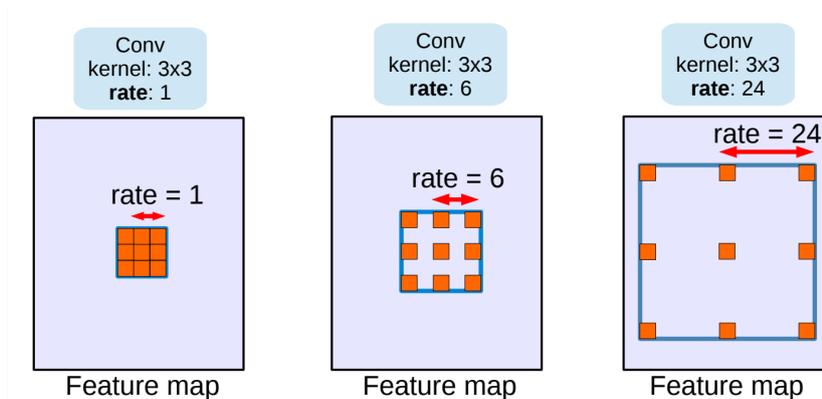


Figura 2.2: Esquemático de la convolución dilatada. Fuente: [26].

rior. Finalmente, se interpola toda la información a distintas escalas en un único mapa de características.

- Estructura de codificador-decodificador: la estructura de la red se puede ver en la imagen 2.1. El codificador se corresponde con la ASPP y se encuentra demarcada por la línea discontinua azul. Los mapas de características resultantes del codificador son interpolados por un factor de 4 y, posteriormente concatenados con características de bajo nivel que tienen la misma resolución. Estas operaciones se corresponden con el decodificador, y están delineadas por la línea discontinua roja.

2.1.2. Modified Aligned Xception

Los autores de Deeplabv3+ modificaron Xception con el objetivo de que el clasificador extrajera mejor las características de las imágenes, lo que permitiría obtener una segmentación más robusta. La red original utiliza la *Deepwise Separable Convolution*, una factorización de la convolución, para disminuir el número de parámetros y el coste computacional de la

red [27].

Las modificaciones que realizaron los autores de Deeplabv3+ fueron tres principalmente [26]: construir una capa oculta más extensa; sustituir las operaciones de *max-pooling* por *Deepwise Separable Convolution*, de manera que se pueda aplicar la convolución dilatada en dichas capas; y por último un mayor uso de *Batch Normalization* y de *ReLU (Rectified Linear Unit)* tras cada *Deepwise Separable Convolution*. Un esquemático de estas modificaciones se puede ver en la Figura 2.3. Los cambios respecto de la red original se encuentran destacados en color naranja. De esta manera, se consigue una mejora decrementando ligeramente el número de parámetros de la red debido a la inclusión de la *trous convolution*.

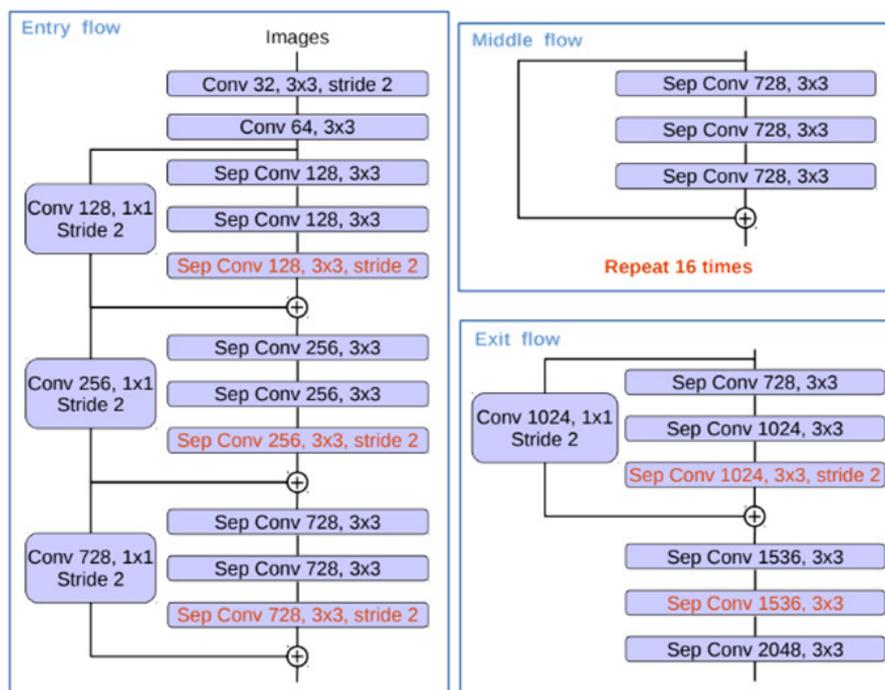


Figura 2.3: Esquema de *Modified Aligned Xception*. Los cambios realizados por los autores se encuentran en naranja. Fuente: [26].

2.1.3. Función de pérdidas, optimizador y activación de la última capa

Debido a la naturaleza del problema, que trata de segmentación de imágenes, y en la que el modelo deberá elegir entre más de una clase, la activación de última capa del modelo fue *Softmax*. La función de pérdidas escogida es la entropía cruzada. La elección del tipo de entropía cruzada depende de la activación de la última capa del modelo, por lo que la función de pérdidas escogida fue *Categorical Crossentropy*. Por último, y debido a su buen ajuste en diversidad de caos, el optimizador fue Adam (*Adaptative Moment Stimation*) [28].

2.2. Bayesian Search

La búsqueda de los hiperparámetros óptimos del modelo es un paso crítico en el diseño de cualquier modelo de *Machine Learning*. Esto es debido al gran impacto que tienen en las prestaciones del modelo. Esta optimización se realiza habitualmente mediante técnicas que no utilicen el descenso del gradiente, debido a que el espacio de búsqueda está compuesto por decisiones discretas, lo que hace que este no sea diferenciable [29]. Generalmente, se utiliza *Grid Search* para este propósito, que consiste en probar todos los valores posibles para cada hiperparámetro dentro de un rango fijo. También otras versiones de esta son utilizadas, como *Random Search*, que simplemente selecciona valores aleatorios del rango de búsqueda definido con *Grid Search*.

En este trabajo, se utilizará *Bayesian Search* para una dirigir el proceso de optimización de una manera eficiente y eficaz. Este método se basa en el teorema de Bayes y es considerado en varios casos la forma más eficiente de optimizar los hiperparámetros [29] [30] [31] [32] [33]. Las distintas funciones que componen esta técnica son:

2.2.1. Función Objetivo

La función objetivo es aquello que se quiere optimizar. En el caso concreto de optimización de hiperparámetros se utilizaría, por ejemplo, el valor de la función de pérdidas (o cualquier otra métrica equivalente) para el conjunto de validación al final del entrenamiento del modelo, que es distinta para cada conjunto de hiperparámetros [30].

2.2.2. Función subrogada

Es la aproximación bayesiana de la función objetivo que puede ser muestreada de manera eficiente. Es una especie de recopilación de todo lo que conocemos de la función objetivo, es decir, todas las pruebas realizadas anteriormente. Las funciones que se utilizan normalmente para este objetivo son los árboles de decisión o los procesos gaussianos [30].

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (2)$$

$$P(A|B) \propto P(B|A) \cdot P(A) \quad (3)$$

Particularizando:

$$P(of|HY) \propto P(HY|of) \cdot P(of) \quad (4)$$

De esta manera, se tiene una cierta información posterior, que no es una probabilidad concreta en sí al haber descartado el denominador del teorema de Bayes, pero que sí es proporcional a esta y puede utilizarse en el proceso de optimización [32]. En la expresión, $P(of|HY)$ es la función subrogada, $P(HY|of)$ se corresponde con una información a posteriori y $P(of)$ es la función objetivo.

2.2.3. Función de adquisición

Esta función utiliza la información posterior, proporcionada por la función subrogada, para seleccionar la siguiente muestra del espacio de búsqueda [33]. En nuestro caso sería la función que, con toda la información que ha sido obtenida previamente, estima qué muestra puede ser la que optimice más el proceso de entre todo el espacio de búsqueda. Esto lo realiza atendiendo a distintos criterios previamente definidos. Las funciones que habitualmente se utilizan son las de *Probability of Improvement (PI)*, *Expected Improvement (EI)* y *Lower Confidence Bound (LCB)*, y se seleccionan en función de intentar balancear la exploración o la explotación en el espacio de búsqueda [30].

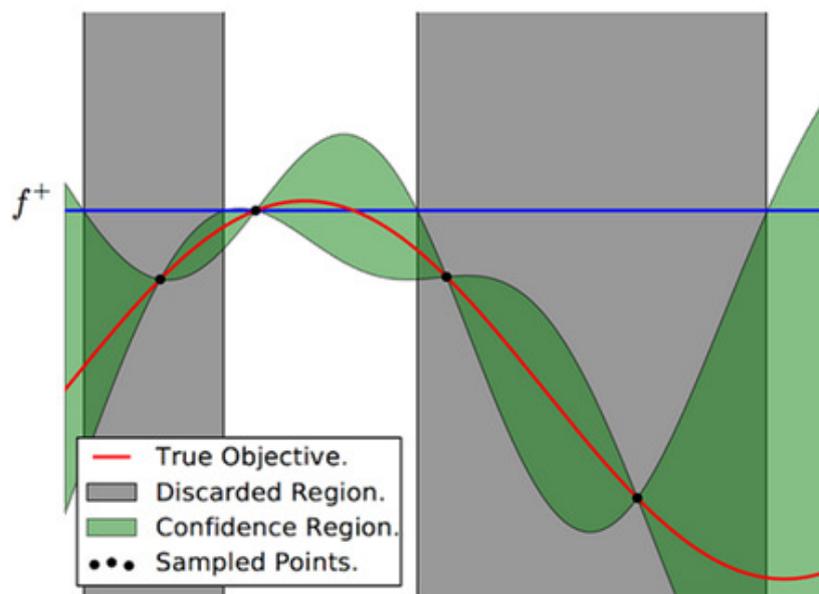


Figura 2.4: Ejemplo función de adquisición. Fuente: [34].

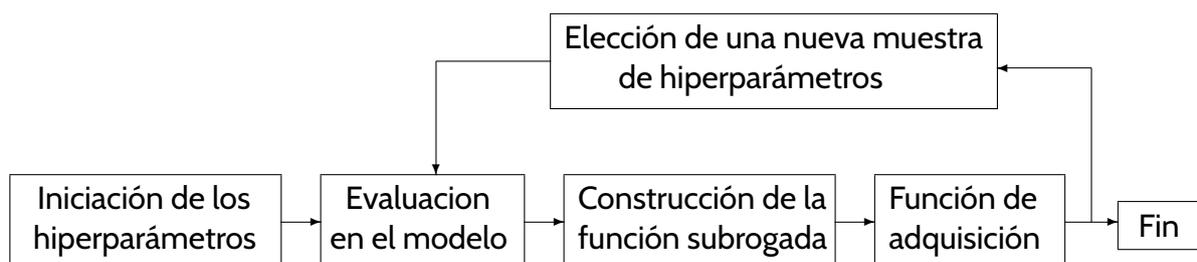


Figura 2.5: Proceso global de *Bayesian Search*. Fuente: propia.

El proceso global se encuentra representado en la Figura 2.5 y comenzaría con una muestra de hiperparámetros, que sería evaluada mediante la función objetivo (el modelo en nuestro caso) y cuyo resultado serviría para actualizar la función subrogada. Finalmente, la función de adquisición seleccionaría el siguiente conjunto de hiperparámetros a evaluar. En la Figura 2.6 se puede apreciar un ejemplo de la evolución del proceso de optimización dependiendo de la función de adquisición elegida. Las Figuras de la primera columna se corresponden con el proceso de optimización mediante *Expected Improvement*, las de la segunda con *Probability of Improvement* y la tercera con *Upper Confidence Bound*. Como se puede apreciar, *PI* opta por explotar el espacio de optimización, mientras que *EI* y *UCB* optan por explorarlo [31]. Un buen balance entre estas funciones es crucial para optimizar adecuadamente el espacio de los hiperparámetros [29].

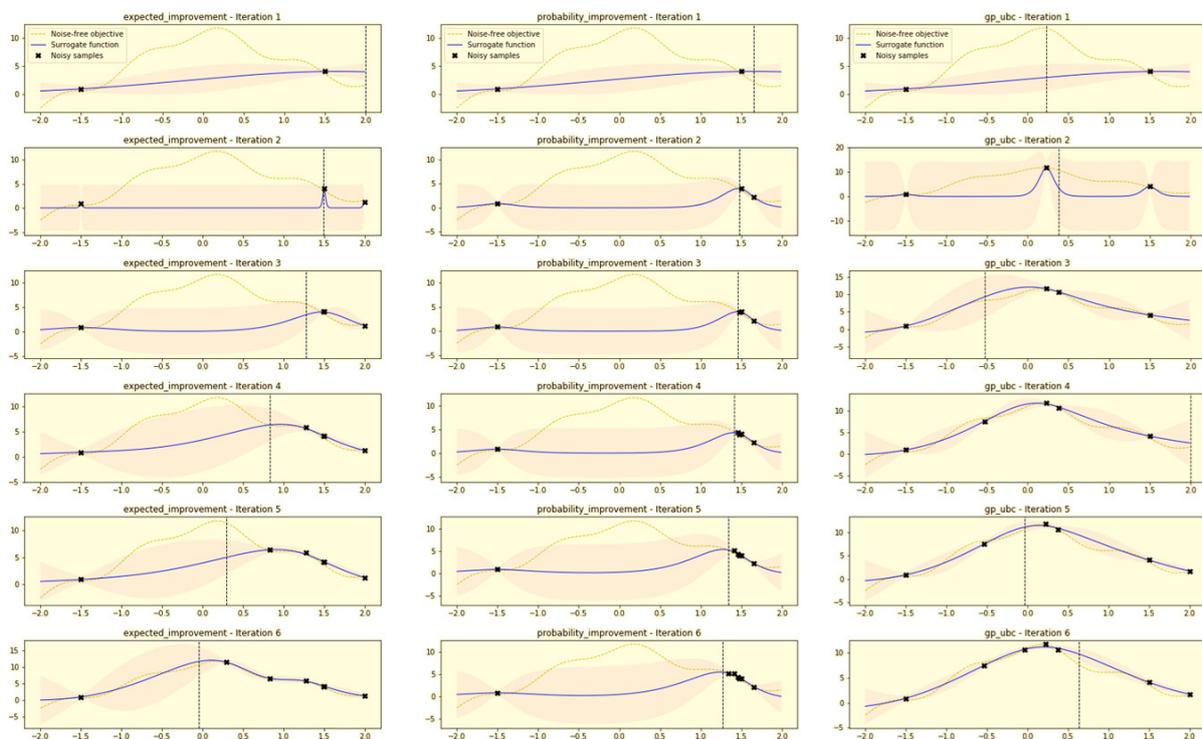


Figura 2.6: Evolución del proceso de optimización en función de la adquisición seleccionada. Fuente: [35].

2.3. Data augmentation

El aumento de datos es una técnica que puede ser bastante beneficiosa a la hora de combatir el sobreentrenamiento del modelo, permitiendo que el modelo generalice mejor [36]. Hoy en día, es un estándar a la hora de diseñar el entrenamiento del modelo, y es extraño encontrar trabajos donde no se utilice.

La idea detrás de esta técnica es realizar cambios aleatorios en las imágenes, de manera

que se obtengan nuevas muestras. Estos cambios deben ser tales que las imágenes, aunque modificadas, puedan entenderse como muestras inalteradas.

Para la modalidad de imágenes con la que estamos trabajando, lo más sensato es aplicar rotaciones, volteos, ajustes de brillo y zooms a las imágenes. Además, también se decidió aplicar deformaciones elásticas y polares, estando esta última específicamente diseñada para imágenes intravasculares, más concretamente en OCT [37].

2.3.1. Deformación elástica

Las deformaciones elásticas son creadas generando, en primer lugar, campos de desplazamiento aleatorio $x(x,y) = rand(-1,1)$ y $y(x,y) = rand(-1,1)$, donde $rand(-1,1)$ es un número aleatorio entre -1 y 1 generado mediante una distribución uniforme. Estos campos son posteriormente convueltos con una distribución gaussiana de desviación estándar σ , que controla la verosimilitud de la transformación [38]. Generalmente, para conseguir resultados más similares a una imagen inalterada, se deben utilizar valores de σ intermedios. En el *paper* original, se recomienda utilizar un valor de $\sigma = 34$ [38].

Por otro lado, para controlar la magnitud del desplazamiento en sí mismo, se utiliza el parámetro α . El comportamiento de la deformación en función de α se puede ver en la Figura 2.7. Cuanto más pequeño sea este valor, menos se controlará el desplazamiento de los píxeles, quedando en un resultado más similar a un ruido gaussiano [38]. Si, por otro lado, se decide tomar un valor demasiado elevado, obtendríamos un resultado más parecido a una traslación.

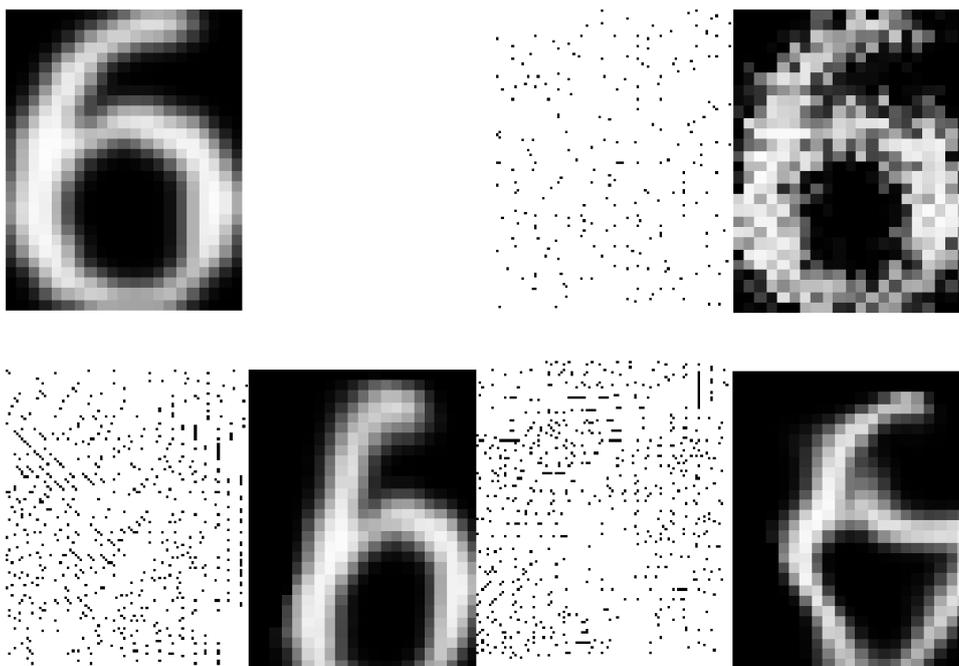


Figura 2.7: Comportamiento de deformación elástica en función de α . Fuente: [38].

2.3.2. Deformación polar

La segunda deformación aplicada es la polar. Esta se utilizó originalmente en imágenes intravasculares de OCT, y fue desarrollada para evitar la deformación de la forma redonda regular del catéter, que no suele ocurrir en la realidad [37]. Por tanto, se trata de una deformación angular y radial en coordenadas polares.

Dada una probabilidad $p_{pd} \sim U(0, 1)$, la imagen deformada I_{def} se construye:

$$I_{def} = \begin{cases} I'_{in}, & \text{if } p_{pd} < 0,5 \\ I'_{in} [F(\theta_i), G(\rho_j)], & \text{otherwise} \end{cases} \quad (5)$$

En esta expresión, (F, G) denota la transformación en el dominio polar, definido por interpolación lineal a trozos en un conjunto de la cuadrícula $\{g_{ij}\}$:

$$\{g_{ij} | (\theta_i, \rho_j), i \in \{0, 1, \dots, N_\theta - 1\}, j \in \{0, 1, \dots, N_\rho\}\} \quad (6)$$

$\{g_{ij}\}$ se define por $\{\theta_i\}$ separando el espacio angular en N_θ sectores cada $2\pi/N_\theta$ y $\{\rho_j\}$ definiendo N_ρ círculos centrados en el catéter. F es la transformación definida con un componente angular:

$$F(\theta_i \pm 2\pi) = \theta_i + \delta\theta_i, \delta\theta_i \sim U\left(-\frac{\pi}{N_\theta}, \frac{\pi}{N_\theta}\right) \text{ for } i \in 0, 1, \dots, N_\theta \quad (7)$$

La continuidad de las transformaciones en todas las direcciones está garantizada por el diseño periódico de F . G es la transformación radial:

$$G(\rho_j) = \rho_j + \delta\rho_j, \delta\rho_j \sim U\left(-\frac{M}{4N_\rho}, \frac{M}{4N_\rho}\right) \text{ for } j \in 2, 3, \dots, N_\rho - 1 \quad (8)$$

El radio $\{\rho_k | k \in 1, 2, N_\rho\}$ se elige de tal manera que el rango de la imagen y el del catéter sufran un cambio mínimo.

Determinando N_θ y N_ρ , el rango de cambio puede modificarse para definir la deformación en una extensión distinta. El rango de ese cambio para $\delta\theta_i$ y $\delta\rho_j$ fue definido usando el número de muestras en ese eje para evitar que se plieguen hacia atrás.

Una representación gráfica de este proceso se puede encontrar en la Figura 2.8.

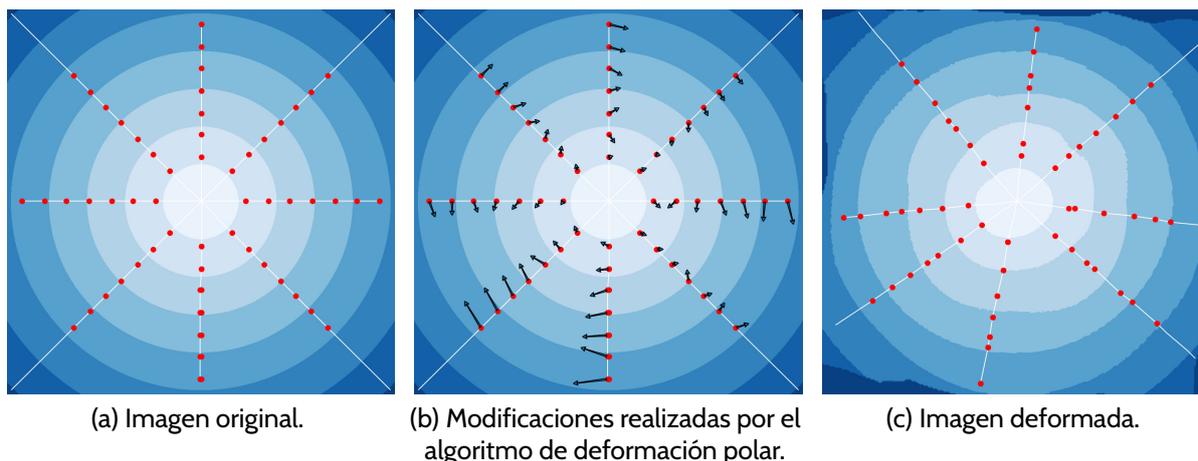


Figura 2.8: Funcionamiento de la deformación polar. Fuente: [37].

2.4. Morfología matemática

La morfología matemática es una herramienta utilizada para extraer los componentes de la imagen que son útiles para la representación y descripción de una región [16]. Se basa en la teoría de conjuntos, de retículos, en topología y en funciones aleatorias. Estas técnicas pueden utilizarse tanto en imágenes binarias como de niveles de gris. En nuestro caso, nos centraremos exclusivamente en el comportamiento en imágenes binarias, debido a que se utilizará en el post-procesado de las predicciones del modelo. Las operaciones contempladas en la morfología matemática son no lineales [16].

Los elementos estructurantes son pequeños conjuntos o subimágenes utilizados para probar propiedades de interés en la imagen bajo estudio [16]. Son el elemento básico a utilizar en este tipo de operaciones. El tipo de elemento estructurante y, obviamente, el tipo de operación, determinarán el resultado final de esta última [16]. En la Figura 2.9 se encuentran ejemplos de distintos elementos estructurantes en color gris.

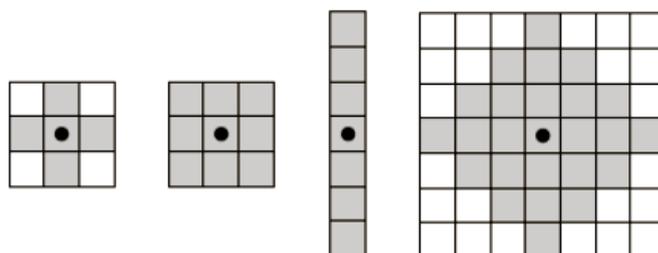


Figura 2.9: Ejemplos de distintos elementos estructurantes. Fuente: [16].

2.4.1. Erosión

La erosión es utilizada para encoger o estrechar los objetos en una imagen binaria [16]. En la Figura 2.10, se puede observar el resultado gráfico de esta operación. La erosión se define de la forma:

$$A \ominus B = \{z | (B)_z \subseteq A\} = \{z | (B)_z \cap A^c = \emptyset\} \quad (9)$$

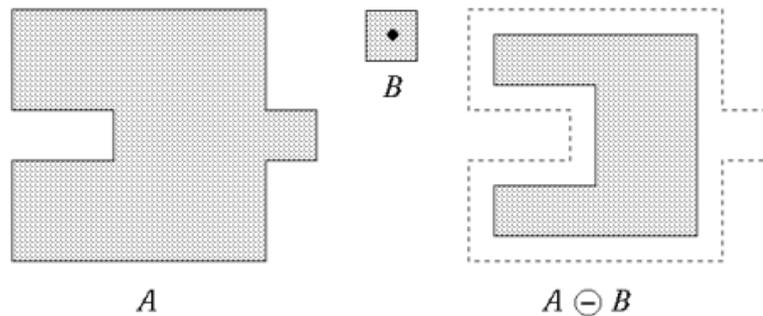


Figura 2.10: Ejemplo de erosión morfológica. Fuente: [16].

2.4.2. Dilatación

La dilatación es utilizada para agrandar los objetos en una imagen binaria [16], y su comportamiento puede verse en la Figura 2.11. Esta operación se define de la forma:

$$A \oplus B = \{z | [(\hat{B})_z \cap A] \subseteq A\} = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (10)$$

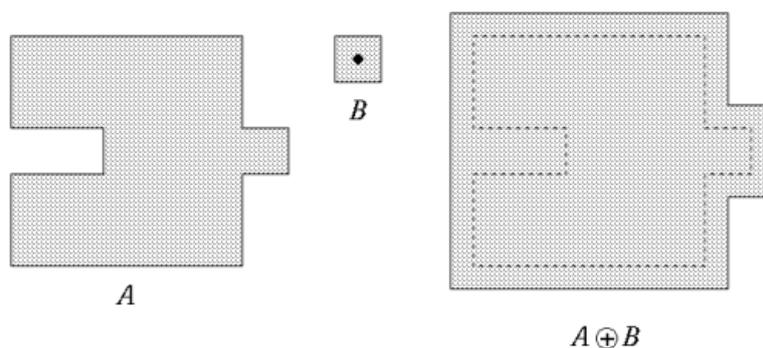


Figura 2.11: Ejemplo de dilatación morfológica. Fuente: [16].

Estas dos operaciones son duales respecto al conjunto complementario y la reflexión cuando el elemento estructurante es simétrico respecto al origen [16]: $(A \ominus B)^c = A^c \oplus \hat{B}$ y $(A \oplus B)^c = A^c \ominus \hat{B}$.

2.4.3. Apertura

La apertura consiste en realizar una erosión seguida de dilatación [16]. Es utilizada para suavizar el contorno del objetivo, y puede entenderse como la eliminación de objetos más pequeños que el elemento estructurante. El comportamiento gráfico de esta operación se puede observar en la Figura 2.12. Se define como:

$$A \circ B = (A \ominus B) \oplus B \quad (11)$$

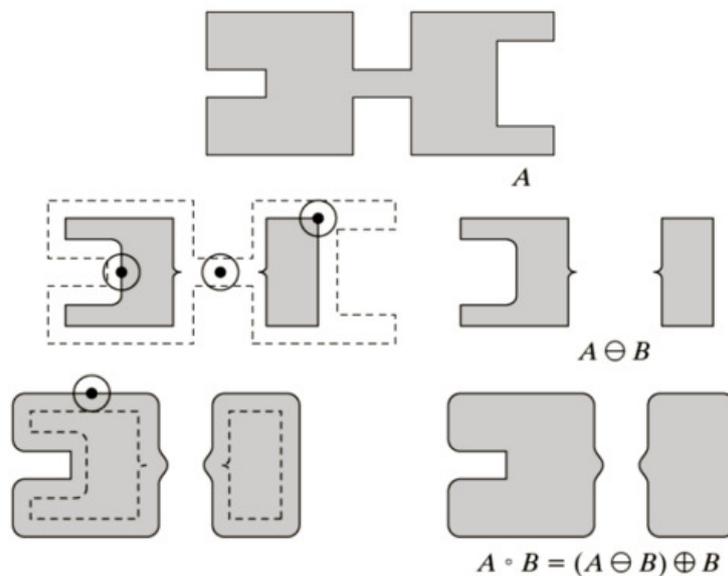


Figura 2.12: Ejemplo de apertura morfológica. Fuente: [16].

2.4.4. Cierre

El cierre se define como dilatación seguida de erosión [16]. Se utiliza para suavizar el contorno de un objetivo, de manera contraria a la apertura, eliminando huecos más pequeños que el elemento estructurante. El resultado de la operación se puede ver en la Figura 2.13. Se define como:

$$A \bullet B = (A \oplus B) \ominus B \quad (12)$$

Al igual que antes, la apertura y el cierre son operadores duales respecto al conjunto complementario y la reflexión cuando el elemento estructurante es simétrico respecto al origen ($\hat{B} = B$): $(A \bullet B)^C = A^C \circ \hat{B}$ y $(A \circ B)^C = A^C \bullet \hat{B}$.

La combinación de apertura y cierre es utilizada principalmente para reducir ruido de distinta modalidad en una imagen, y esto es lo que se centrará este trabajo para intentar

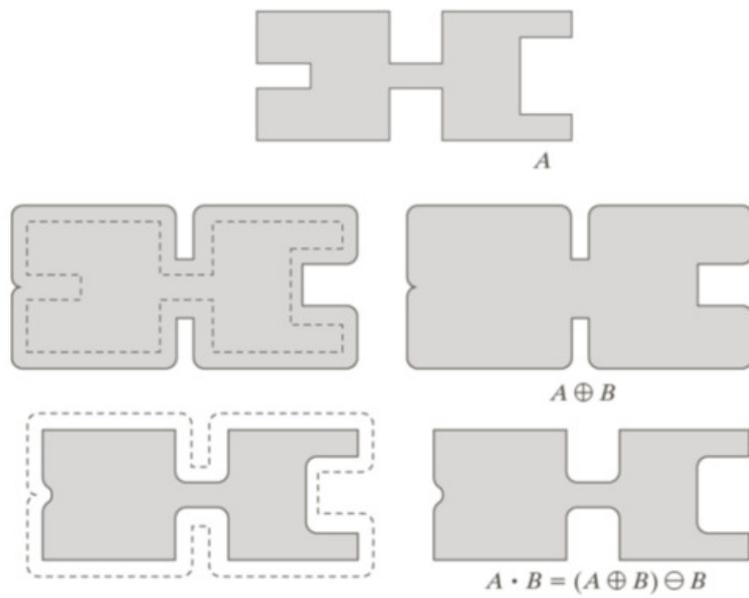


Figura 2.13: Ejemplo de cierre morfológico. Fuente: [16].

unir regiones no aisladas. Otra aplicación es la extracción de contornos de la imagen, cosa que también se hará en este trabajo.

3 Materiales y métricas

3.1. Base de Datos

Los datos utilizados fueron adquiridos en el *FFR* (Fractional Flow Reserve) *react trial* [39], en el que participaron 290 pacientes. En este trabajo se pretendía evaluar si los pacientes, tras la PCI guiada por IVUS y con un *FFR* menor de 0.90, es decir, una relación entre el flujo máximo posible de sangre en el vaso sanguíneo con y sin obstrucción mayor que 0.90 [39], podían mejorar la tasa de fallo del *stent* en un tiempo de 1 año respecto a los no guiados por IVUS.

Para ello, el cardiólogo analizaba el *pullback* de cada paciente *frame a frame* tanto antes como después de la intervención [39]. Se realizaba el etiquetado del lumen y la pared vascular, y se calculaba automáticamente el *FFR*. Aquellos pacientes que obtuvieran un *FFR* mayor que 0.90 se consideraron para este estudio.

En este Trabajo de Fin de Estudios, se tomaron los *pullbacks* que conforman cada una de las distintas adquisiciones. Las segmentaciones están realizadas cada 12 *frames* en el *pullback* y, en caso de que la pared vascular no sea visible en 180 grados, el cardiólogo no realiza dicha segmentación manual [4]. Esto provoca que haya casos en los que tengamos segmentación del lumen pero no de la pared vascular. Estos casos son directamente descartados y son 23 *frames* en total. Se dispone de 8270 imágenes sin *stent* y 3628 con *stent*, lo que conforma un total de 11898 imágenes, con una resolución de 512×512 .

3.2. Funciones de I/O

La forma en la que estaban guardados los datos en memoria permitía su identificación unívoca dentro del *pullback*. En primer lugar, se tenían en distintos directorios tanto las entradas del modelo sin segmentar (el *pullback*), como las etiquetas (las segmentaciones de los cardiólogos). En el caso de los *frames* sin segmentar, se encontraba cada *pullback* numerado y dos archivos distintos; el RAW y la cabecera MHD.

En caso de las etiquetas, estas se encontraban dentro de distintas carpetas: lumen,

stent y *vessel*. Las etiquetas se almacenaron en formato PNG y se identificaban con el nombre correspondiente a la posición que ocupan dentro del *pullback*. Se utilizó esta información para unir la etiqueta del lumen y la pared vascular correspondientes a un mismo *frame* en una única etiqueta mediante la suma. Además, se cargaba el archivo de *pullbacks* completo y se descartaban aquellas etiquetas del lumen que no tienen pared vascular. Finalmente, estas etiquetas se codificaban en *one-hot*. El proceso global se encuentra esquematizado en la Figura 3.1

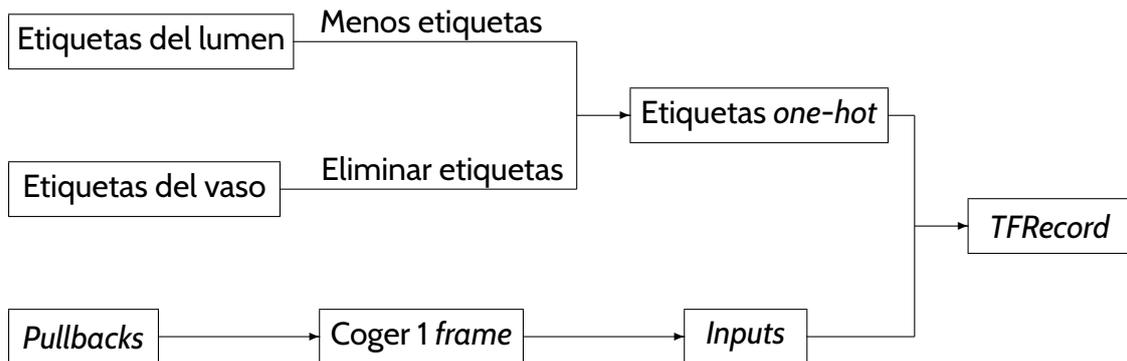


Figura 3.1: Flujograma del manejo de los datos. Fuente: propia.

Por último, se tomaban los *inputs* y las etiquetas y se codificaban como una tupla en un archivo **TFRECORD**, que posteriormente alimentaba al modelo de Tensorflow [40] y que permitía entrenar al mismo sin cargar datos en memoria RAM (*Random Access Memory*), evitando el *overflow* debido a la baja RAM de la GPU (*Graphics Processing Unit*), y al elevado número de muestras disponibles.

3.2.1. Manejo de *data.Dataset*

La función `tf.data.TFRecordDataset` permite cargar el archivo **TFRECORD** anteriormente descrito para ser convertido en un objeto `data.Dataset`, que será el que pasaremos a nuestro modelo para entrenar o inferir.

Este objeto permite mapear una serie de funciones a la hora de entrenar el modelo con funcionalidades interesantes. En nuestro caso, se utiliza `shuffle`, que baraja los datos antes de cada época, y `batch`, que permite cargarlos en conjuntos de muestras. La opción `prefetch` con `AUTOTUNE` permite asignar todos los recursos *hardware* disponibles para que el modelo los utilice de la manera mas óptima posible.

Por último, las técnicas de aumento de datos utilizadas en este caso están implementadas gracias a una función que las incluye a todas. En nuestro caso, recordemos que se define un parámetro que controla la probabilidad de que una imagen sea modificada. Las alteraciones aleatorias aplicadas para cada imagen son las previamente definidas en el capítulo anterior. Estas son rotaciones, volteos, cortes, *zooms*, ajustes de brillo y una de las dos

deformaciones expuestas.

3.3. Metodología

Inicialmente, se realizó una familiarización con las imágenes de IVUS: el tipo de estructuras que se pueden encontrar, la manera de diferenciar la pared vascular del lumen y otras modalidades de imágenes que pueden encontrarse. Todo esto se hizo explorando toda la base de datos disponible y construyendo distintas funciones mediante el *software* de [MeVisLab](#), que permite integrar funciones propias en código de Python [41] y simplificar la tarea de visualización de imágenes.

Posteriormente, se decidió investigar sobre los distintos trabajos previamente realizados en cuanto a segmentación de pared vascular y lumen se refiere, para comprender el *state-of-the-art* del campo y ver las posibles técnicas que pudieran aplicarse al algoritmo.

Como siguiente paso, se comenzaron a escribir las funciones de I/O del modelo hasta seguir la estructura de datos de [TFRECORD](#), para después hacer lo propio con el modelo con Tensorflow y comenzar los primeros entrenamientos. En este punto, se descubrió uno de los problemas que tenía utilizar objetos [data.Dataset](#) con la arquitectura específica de Deeplabv3+ [26]: la capa *Batch Normalization* posterior a la ASPP provocaba un error numérico de *null*, lo que provocaba errores durante el entrenamiento del modelo. Esto se pudo resolver de manera directa cambiando la implementación en Tensorflow [40] a una que hacía un 15 % más lento el procesado y permitía entrenar el modelo de igual manera, sin errores.

Una vez el modelo estuvo listo para entrenar, se buscaron los hiperparámetros óptimos mediante *Bayesian Search*, y se obtuvo el cómputo del modelo. Finalmente, se buscó mejorar los resultados debido a la alta Distancia Hausdorff en comparación con otros modelos del *state-of-the-art* mediante técnicas de morfología matemática. Se emplearon aperturas y cierres con elementos estructurantes de tamaño progresivo, consiguiendo mejorar las prestaciones considerablemente, especialmente en imágenes con bifurcaciones.

3.4. Métricas

Etiquetas a evaluar

En el caso de este trabajo, se ha decidido evaluar la bondad del modelo tomando tres etiquetas distintas: el lumen, la pared vascular y la Región de Interés (Region Of Interest, ROI), que se definió como la región correspondiente a la pared vascular que no se correspon-

de con el lumen. La ROI consigue condensar de cierta manera las prestaciones del modelo en un único número, ya que los errores cometidos por el modelo en asignación de píxeles al lumen o la pared vascular tienen mayor impacto en esta etiqueta.

Área Bajo la Curva

El **AUC** (*Area Under the Curve*) es una métrica relacionada con la precisión de clasificación de píxel (en caso de trabajar con imágenes), y computa el área bajo la curva discretizada de los valores de la *precision* sobre el *recall*. Esto se denomina Curva característica Operativa del Receptor (o ROC), e ilustra la habilidad del clasificador cuando se varía el umbral de discriminación. El rango de esta métrica se sitúa entre 0 y 1. Si el valor es cercano a 0.5, se considerará que el clasificador está actuando de manera aleatoria, por lo que no estará funcionando de manera correcta. Se utilizará la implementación de Scikit-Learn [42] para obtener el **AUC**.

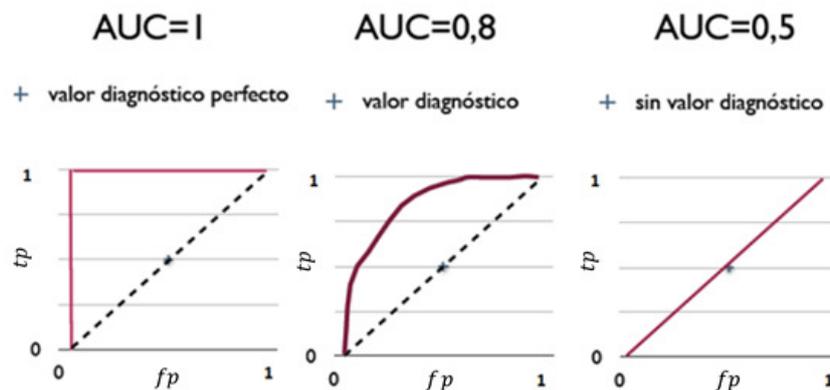


Figura 3.2: Conclusiones extraíbles de los valores de AUC. Fuente: [43].

Índices de Jaccard y Dice

En este caso, estos dos estadísticos se centran en medir la similitud entre dos muestras, a través de la unión e intersección de conjuntos [44]. En primer lugar, el **Índice de Jaccard** (JI) o *Intersection over Union* (IoU) es el cociente entre la intersección y la unión de dos conjuntos de datos [44], y se define de la forma:

$$JI(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} = \frac{tp}{tp + fn + fp} \quad (1)$$

Por otro lado, el Índice de Dice (DSC), que se puede obtener a partir del **Índice de Jaccard**, intenta realizar esta medida de una manera distinta:

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|} = \frac{2tp}{(tp + fp) + (tp + fn)} = \frac{2J(A, B)}{J(A, B) + 1} \quad (2)$$

Estos dos índices tomarán valores entre 0 y 1, y serán mejores cuanto más próximos sean a 1, ya que las dos segmentaciones tendrán más superficie en común.

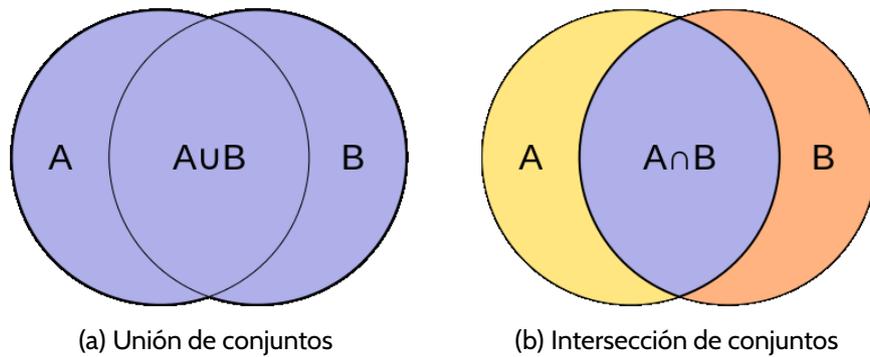


Figura 3.3: Intersección y unión de conjuntos. Fuente: [44].

El **Índice de Jaccard** fue calculado utilizando la librería de Scikit-learn [42], y el Índice de Dice fue obtenido mediante la expresión 2.

Distancia Hausdorff

La Distancia Hausdorff mide cuán lejos se encuentran dos subconjuntos de un espacio métrico [45]. Dos conjuntos están próximos si todo punto del primer conjunto está próximo a algún punto del segundo conjunto, y viceversa. La definición formal es la siguiente:

$$d_H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right\} \quad (3)$$

En la Figura 3.4 se muestra un ejemplo gráfico del cálculo de la **Distancia Hausdorff** entre dos figuras. A la hora de interpretar los resultados numéricos, interesará que esta medida tome los valores más bajos posibles, ya que se considerarán mejores predicciones.

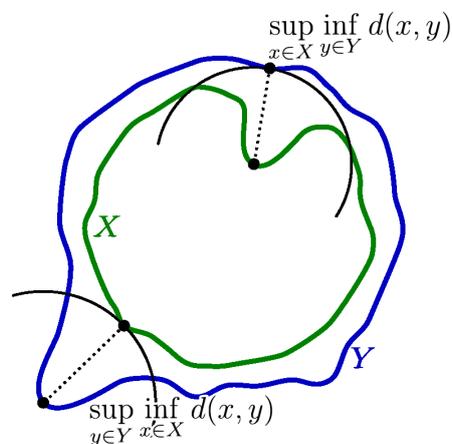


Figura 3.4: Ejemplo de medida de Distancia Hausdorff. Fuente: [45].

En el caso de este trabajo, se convertirá esta **Distancia Hausdorff** de píxeles a milímetros, gracias a la relación en el *pullback* de $0,015625 \text{ mm/pixel}$. Además, para agilizar el cómputo de esta métrica, que generalmente es de las más pesadas, se decidió extraer el contorno de las muestras mediante erosión. La implementación de **Distancia Hausdorff** utilizada fue la proporcionada por Scipy [46].

4 Resultados

4.1. Vecindario

Inspirado por otros trabajos, se decidió utilizar *frames* contiguos para construir imágenes multiespectrales de tres canales, de manera que se ayude en la segmentación al haber añadido más información. Estas imágenes se construían definiendo un vecindario de 11 imágenes alrededor del *frame*, tomando las otras muestras disponibles del *pullback* correspondiente. De estas imágenes se seleccionan dos atendiendo a un criterio concreto, y se añaden como nuevos canales a la imagen en niveles de gris. Es importante recordar que esto se puede hacer gracias a los cardiólogos que etiquetaron la base de datos con una distancia de 12 *frames*.

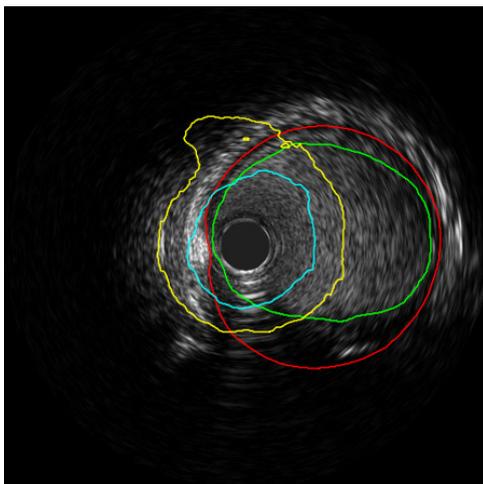
Los criterios para la elección de las imágenes se basan principalmente en correlación, y son los siguientes: 1) utilizar una única imagen; 2) elegir las dos imágenes con mayor correlación en un vecindario de 11 *frames* para formar una entrada de tres canales; 3) elegir las dos imágenes con menor correlación en un vecindario de 11 *frames* para formar una entrada de tres canales; 4) elegir las dos imágenes adyacentes para formar una entrada de tres canales. Se hizo una comparación de cada uno de estos métodos con no utilizar ninguna información del vecindario.

Métrica	1 <i>frame</i>	3 más	3 menos	3 ady
Lumen				
AUC	0,961 ± 0,009	0,808 ± 0,014	0,807 ± 0,016	0,798 ± 0,013
Jl	0,881 ± 0,017	0,501 ± 0,018	0,506 ± 0,017	0,494 ± 0,013
DSC	0,937 ± 0,033	0,668 ± 0,035	0,672 ± 0,033	0,661 ± 0,026
HD (mm)	0,453 ± 0,068	1,246 ± 0,057	1,208 ± 0,042	1,328 ± 0,033
Pared vascular				
AUC	0,960 ± 0,010	0,693 ± 0,018	0,707 ± 0,020	0,713 ± 0,021
Jl	0,892 ± 0,023	0,38 ± 0,039	0,412 ± 0,040	0,396 ± 0,031
DSC	0,937 ± 0,014	0,531 ± 0,049	0,572 ± 0,043	0,557 ± 0,029
HD (mm)	0,997 ± 0,068	1,818 ± 0,087	1,749 ± 0,090	1,862 ± 0,083

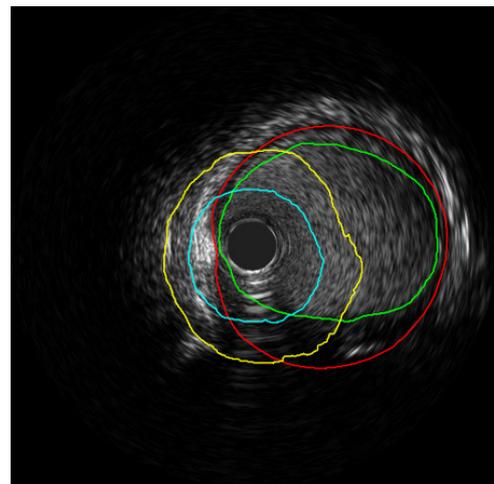
Tabla 4.1: Métricas de los cuatro enfoques seleccionados.

En la Tabla 4.1 se pueden ver los resultados correspondientes a la segmentación del lumen y la pared vascular para los enfoques seleccionados. Para este experimento, se tomaron unos 10 *pullbacks* de los 58 disponibles, y se realizaron 3 *fold*s, por lo que la desviación estándar mostrada es entre *fold*s.

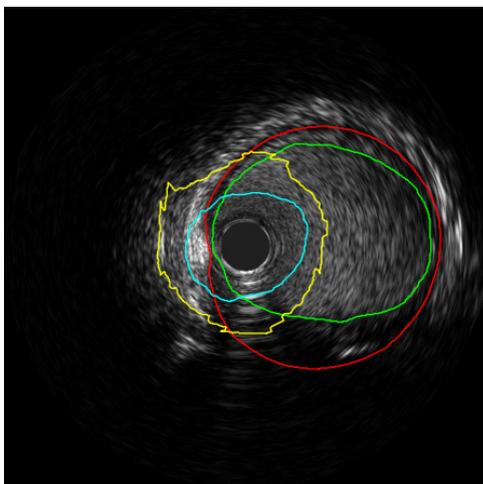
Tanto en la media como en dicha desviación estándar se aprecia que el hecho de utilizar esta información del vecindario no ayuda a mejorar las prestaciones del modelo, siendo el mejor resultado el correspondiente al primer enfoque, es decir, utilizar el *frame* correspondiente a la etiqueta en cuestión.



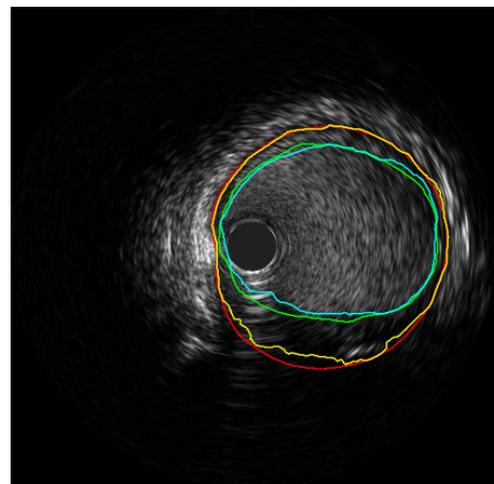
(a) Resultado al tomar las imágenes más parecidas del vecindario.



(b) Resultado al tomar las imágenes menos parecidas del vecindario.



(c) Resultado al tomar las imágenes adyacentes.



(d) Resultado al tomar la imagen en escala de grises.

Figura 4.1: Contornos de la predicción al utilizar una única imagen e imágenes multispectrales de 3 canales como *input* del modelo.

Por otro lado, en la Figura 4.1 se tienen ejemplos gráficos de las segmentaciones, encontrando resultados muy erráticos en los casos de entrenar el modelo con tres muestras en lugar de una. Las líneas en color lima y cian representan la etiqueta y la predicción del

lumen, respectivamente. De igual manera, las líneas roja y amarilla se corresponden con la etiqueta y predicción de la pared vascular, respectivamente.

4.2. Selección de los hiperparámetros con *Bayesian Search*

La utilización de métodos bayesianos para la optimización de los hiperparámetros de los modelos es una de las cuestiones abordadas en este trabajo. Como ya se presentó en el apartado 2.2, distintos trabajos muestran el beneficio en cuanto a prestaciones y computación de utilizar este tipo de algoritmos en el proceso de optimización de los hiperparámetros del modelo.

Los hiperparámetros del modelo son aquellos parámetros que controlan el tamaño del modelo, como *OS* y *Middle flow size*, y los parámetros del optimizador, como *Learning Rate*, β_1 y β_2 . También se optimizó el parámetro que controla la probabilidad de modificar una imagen, que fue denominado DA_{th} . En la Tabla 4.5 se tiene una comparativa de las prestaciones del modelo previo a la búsqueda de los hiperparámetros óptimos y tras la realización de la misma.

Los hiperparámetros utilizados en primera instancia fueron: $OS = 16$, *middle flow size* = 16, *learning rate* = 0,001, $\beta_1 = 0,9$, $\beta_2 = 0,999$ y $DA_{th} = 0,75$. Por su parte, el *batch size* se eligió de manera automática para que fuera el máximo posible que permitiera al modelo caber en la memoria RAM.

El resultado de la optimización, tras 49h de cómputo, proporcionó los siguientes valores: $OS = 8$, *middle flow size* = 12, *learning rate* = 0,002, $\beta_1 = 0,810$, $\beta_2 = 0,940$ y $DA_{th} = 0,9$. En el caso de este modelo, el máximo *batch size* que evitaba el *overflow* durante el entrenamiento del modelo fue 4.

	Métricas	Lumen	Pared vascular	ROI
Antes	AUC	0,963 ± 0,008	0,960 ± 0,010	0,882 ± 0,026
	Jl	0,882 ± 0,015	0,892 ± 0,019	0,665 ± 0,048
	DSC	0,936 ± 0,009	0,941 ± 0,011	0,787 ± 0,038
	HD (mm)	0,292 ± 0,045	0,650 ± 0,044	0,391 ± 0,053
Después	AUC	0,974 ± 0,002	0.974 ± 0.005	0,913 ± 0,011
	Jl	0,909 ± 0,005	0,925 ± 0,007	0,737 ± 0,017
	DSC	0,952 ± 0,003	0,960 ± 0,004	0,839 ± 0,013
	HD (mm)	0.210 ± 0.010	0.475 ± 0.056	0.292 ± 0.023

Tabla 4.2: Métricas antes y después de la optimización bayesiana.

Se puede observar en los resultados que tanto la media como la desviación estándar entre *folds* mejora tras la optimización bayesiana. Aunque haya mejoría en todas las métricas tras el entrenamiento del modelo, la métrica que mayor mejora obtiene es la Distancia

Hausdorff, siendo de un 26,9 % en la pared vascular, de un 28,1% en el lumen y de un 25,3 % en la ROI. En cuanto al resto de métricas, es también notable la mejoría de la AUC de la pared vascular, siendo el resto de unas pocas centésimas y, en algunos casos, de milésimas.

Elección de deformación

En la sección 2.3 se presentó las dos deformaciones que se añadirían a las técnicas de aumento de datos de este trabajo. En este apartado, se pretende estudiar cuál de estas dos deformaciones tiene un mejor impacto en el aprendizaje del modelo: la deformación polar o la elástica, de manera que se utilice exclusivamente una de ellas. Esto es debido a que la deformación polar es una extensión de la elástica, y se entiende que realizan funciones similares, de manera que unidas únicamente añadirían ruido indeseado al entrenamiento del modelo.

En primer lugar, se utilizó de nuevo la optimización bayesiana de manera aislada para los parámetros de estas deformaciones. Los resultados de la deformación elástica fueron $\alpha = 34$ y $\sigma = 4$, y los de la deformación polar fueron $N_{angle} = 75$, $N_{\rho} = 15$, $C_{angle} = 0,13$, y $C_{\rho} = 0,05$.

Posteriormente, se entrenaron los modelos bajo las mismas condiciones haciendo 5 *fold*s, para comprobar cuál de las deformaciones funcionaba mejor. La media de los resultados correspondientes se encuentran en la Tabla 4.3, e indican que la deformación polar mejora el aprendizaje del modelo frente a la elástica.

Def	Etiqueta	AUC	Jl	DSC	HD (mm)
Elástica	Lumen	0,978 ± 0,019	0,669 ± 0,116	0,796 ± 0,085	0,871 ± 0,592
	Vessel	0,971 ± 0,022	0,756 ± 0,154	0,852 ± 0,108	0,646 ± 0,440
	ROI	0,696 ± 0,130	0,468 ± 0,312	0,638 ± 0,476	0,311 ± 0,522
Polar	Lumen	0,978 ± 0,019	0,773 ± 0,104	0,868 ± 0,069	0.545 ± 0.280
	Vessel	0,980 ± 0,017	0,809 ± 0,128	0,889 ± 0,084	0,541 ± 0,349
	ROI	0.816 ± 0.112	0.321 ± 0.191	0.486 ± 0.320	0,432 ± 0,309

Tabla 4.3: Comparación entre las deformaciones elástica y polar.

En este caso, la diferencia principal se encuentra en la ROI, marcando que el área de interés a ser segmentado tiene mejores métricas en cuanto a área y AUC se trata. Por otra parte, la mejoría correspondiente a la Distancia Hausdorff del lumen es bastante destacable (en negrita).

4.3. Postprocesado

Varias predicciones presentaban contornos irregulares y erráticos, por lo que se decidió utilizar operaciones de morfología matemática para tratar de conseguir regularizarlos, eliminando el ruido de alta frecuencia. Se esperaba que la mejora de estos contornos tuviera un impacto directo positivo en la Distancia Hausdorff.

Algunos ejemplos de las segmentaciones sin postprocesar se pueden visualizar en la Figura 4.2. Para intentar resolver este problema, se utilizaron aperturas y cierres con elementos estructurantes de distinto tamaño de manera progresiva, en función del tamaño de la etiqueta de la pared vascular. Se encontró que la forma de elemento estructurante que mejor resultado daba era el rombo.

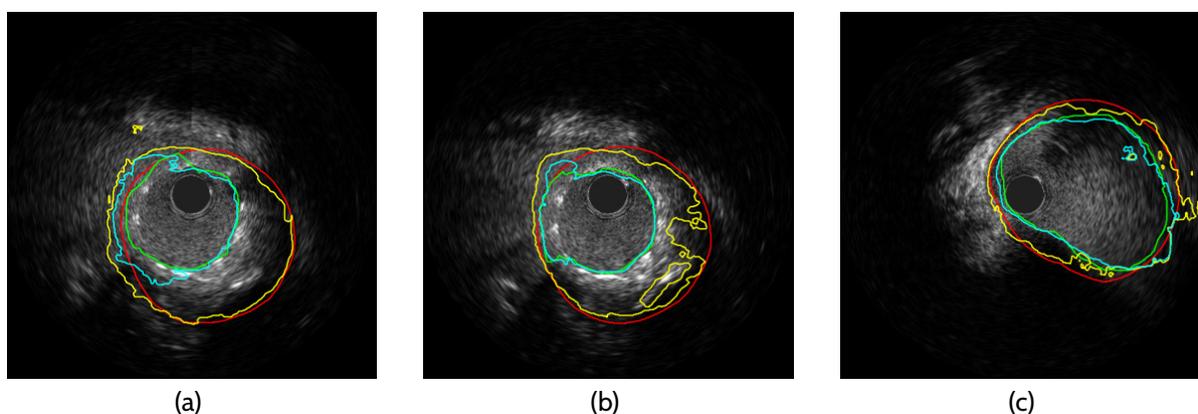


Figura 4.2: Imágenes con sobresegmentación y contornos irregulares sin postprocesar.

Los resultados detallados de cada *fold*, así como la media de todos los resultados, con la desviación estándar tanto entre imágenes como entre *folds*, se puede ver en la Tabla 4.4. Los resultados muestran una mejora bastante notable, especialmente en la Distancia Hausdorff de la pared vascular. Como referencia para la mejora de este postprocesado, se pueden tomar la Tabla 4.2 y la media que aparece al final de la Tabla 4.4, mostrando una mejoría notable en todas las métricas de la pared vascular y en la ROI.

Se puede apreciar que los resultados son bastante parejos para los distintos *folds*, siendo el *fold* 3 el que mejores resultados obtiene en todas las métricas. La baja desviación estándar de la fila correspondiente a la media (en negrita) marca la poca diferencia entre *folds*, mientras que la desviación estándar del *fold* 3 marca la menor diferencia entre segmentaciones para ese *fold*, prueba de que el entrenamiento ha aprendido ligeramente mejor en ese caso.

Por otra parte, en la ROI se puede apreciar las mejores prestaciones del *fold* 3 respecto al resto de *folds* con mejor detalle, destacando la Distancia Hausdorff y el Índice de Dice para

	Métrica	Lumen	Pared vascular	ROI
Fold 1	AUC	0,977 ± 0,018	0,975 ± 0,025	0,910 ± 0,064
	Jl	0,910 ± 0,054	0,929 ± 0,053	0,687 ± 0,242
	DSC	0,952 ± 0,032	0,963 ± 0,031	0,780 ± 0,249
	HD (mm)	0,207 ± 0,157	0,266 ± 0,206	0,231 ± 0,171
Fold 2	AUC	0,972 ± 0,023	0,974 ± 0,030	0,917 ± 0,062
	Jl	0,906 ± 0,060	0,922 ± 0,064	0,694 ± 0,228
	DSC	0,950 ± 0,037	0,958 ± 0,038	0,789 ± 0,230
	HD (mm)	0,219 ± 0,240	0,280 ± 0,267	0,247 ± 0,184
Fold 3	AUC	0,976 ± 0,017	0,981 ± 0,017	0,931 ± 0,049
	Jl	0,918 ± 0,046	0,936 ± 0,046	0,730 ± 0,213
	DSC	0,957 ± 0,027	0,966 ± 0,026	0.818 ± 0.214
	HD (mm)	0,193 ± 0,217	0,234 ± 0,222	0.214 ± 0.155
Fold 4	AUC	0,972 ± 0,021	0,970 ± 0,030	0,901 ± 0,070
	Jl	0,904 ± 0,055	0,921 ± 0,061	0,680 ± 0,233
	DSC	0,949 ± 0,032	0,958 ± 0,036	0,777 ± 0,236
	HD (mm)	0,218 ± 0,190	0,303 ± 0,251	0,244 ± 0,172
Fold 5	AUC	0,972 ± 0,023	0,970 ± 0,034	0,909 ± 0,067
	Jl	0,909 ± 0,053	0,919 ± 0,066	0,677 ± 0,245
	DSC	0,951 ± 0,032	0,956 ± 0,040	0,771 ± 0,254
	HD (mm)	0,210 ± 0,182	0,309 ± 0,254	0,251 ± 0,191
Media	AUC	0,974 ± 0,002	0,974 ± 0,005	0,914 ± 0,011
	Jl	0,909 ± 0,005	0,925 ± 0,007	0,694 ± 0,021
	DSC	0,952 ± 0,003	0,960 ± 0,004	0,787 ± 0,019
	HD (mm)	0,207 ± 0,015	0,278 ± 0,030	0,237 ± 0,016

Tabla 4.4: Métricas de tres etiquetas distintas con datos sin *stent*.

dicho *fold*. De manera general, las métricas aquí presentadas son bastante buenas, dando lugar a segmentaciones muy precisas desde el punto de vista de todas las métricas.

En la Figura 4.3, se presenta una distribución en forma de diagrama de violín con las métricas correspondientes al lumen, la pared vascular y la Región de Interés para el *fold 3*. Este diagrama contiene la forma de distribución de los datos, representando de esta manera la frecuencia de un determinado valor dentro de los datos en el eje Y y su valor en el eje X, permitiendo comprobar cuán dispersos están los valores.

Como era esperado, las métricas correspondientes a la ROI no son tan consistentes como para el lumen o la pared vascular, teniendo una distribución de la ROI bastante extendida, mientras que las correspondientes al lumen y a la pared vascular están considerablemente comprimidas.

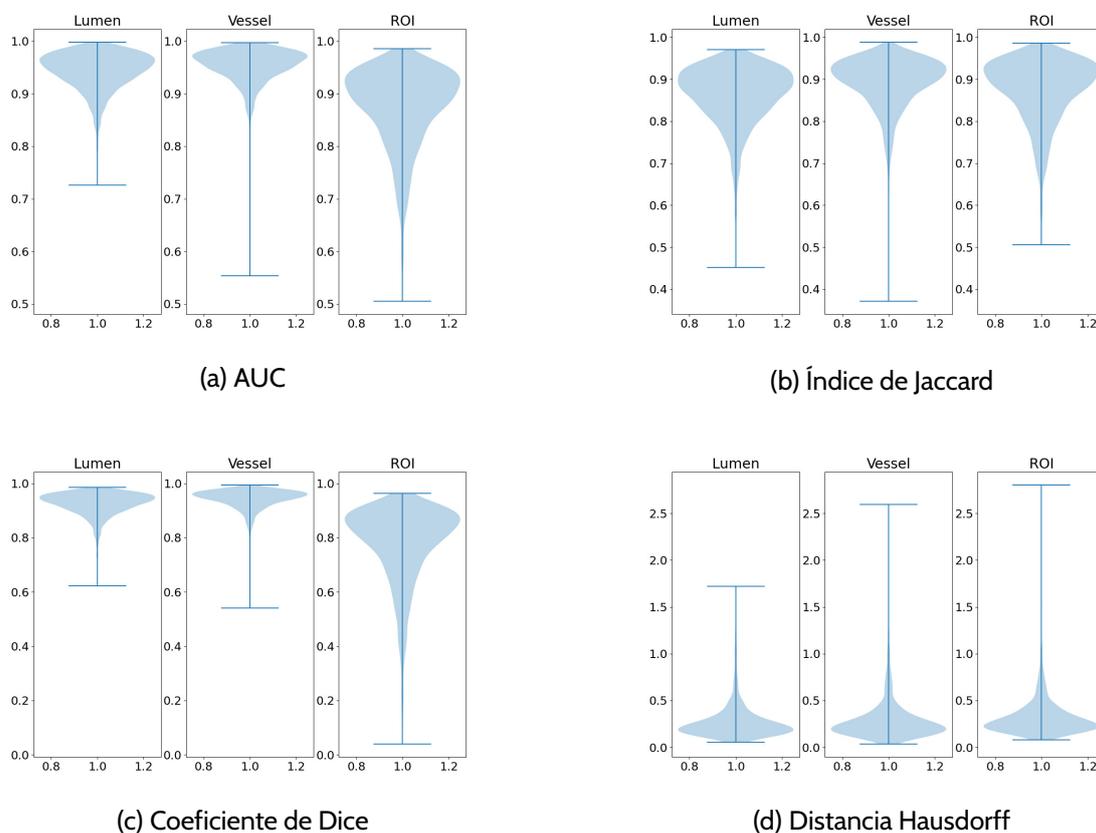


Figura 4.3: Diagramas de violín para el *fold 3* de las distintas métricas.

La Figura 4.3 muestra que los diagramas de violín correspondientes al Índice de Jaccard tienen mayor dispersión. En general, todas las métricas siguen una distribución similar, teniendo gran parte de las muestras valores muy buenos para las métricas y descendiendo progresivamente el valor de las métricas.

En la Figura 4.4, se muestran varios ejemplos de contornos detectados sobre imágenes de IVUS. En la primera fila, se puede observar casos en los que el modelo ha tenido

una predicción errática, principalmente en bifurcaciones debido a que tanto el lumen como la pared vascular presentan una forma ovalada y más alargada que en casos normales. Por otra parte, se puede encontrar una segunda fila conteniendo las mejores predicciones del modelo para dicho *fold*, que suelen presentar una forma circular y regular. Las líneas de color lima y cian se corresponden con la etiqueta del lumen y la predicción del lumen, respectivamente. De igual manera, la línea roja se corresponde con la etiqueta de la pared vascular, y la predicción del modelo se puede ver en amarillo.

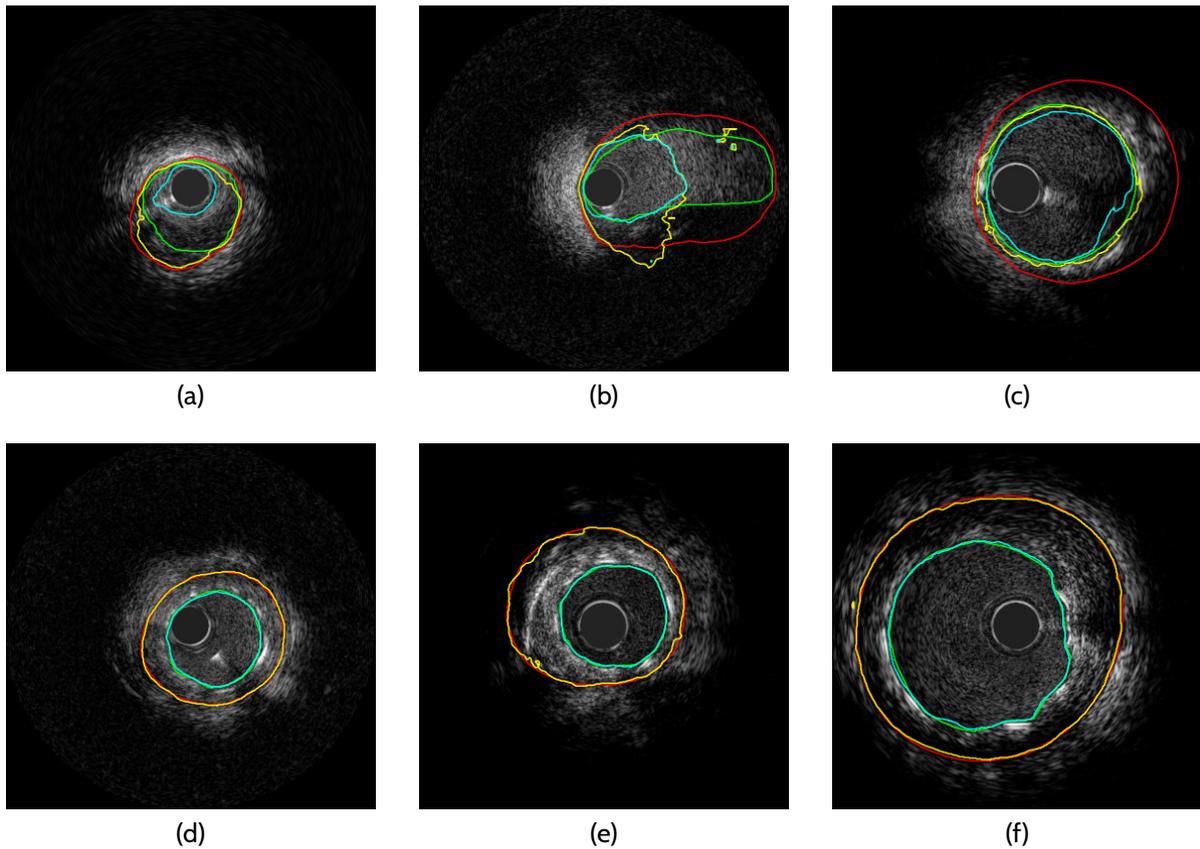


Figura 4.4: Contornos de predicciones sin *stent* del modelo final optimizado.

4.4. *Stent*

Debido a que se tenían las etiquetas correspondientes al *stent*, también se aplicó este flujo de trabajo a este caso. Se encontró un resultado satisfactorio en la segmentación, especialmente en el lumen, como se puede ver en los resultados representados en la Tabla 4.5. Se observa una diferencia mayor entre prestaciones de *fold*s, pero aún así se obtiene un valor medio bastante elevado para el Índice Dice y una Distancia Hausdorff bastante baja para el lumen.

En este caso, se puede ver que hay una mayor diferencia en las prestaciones entre distintos *fold*s, principalmente en la media de la Distancia Hausdorff de todas las etiquetas

	Métrica	Lumen	Pared vascular	ROI
Fold 1	AUC	0,980 ± 0,013	0,948 ± 0,046	0,884 ± 0,083
	Jl	0,936 ± 0,038	0,879 ± 0,089	0,669 ± 0,135
	DSC	0,967 ± 0,022	0,933 ± 0,054	0,802 ± 0,237
	HD (mm)	0,424 ± 0,795	0,571 ± 0,652	0,374 ± 0,366
Fold 2	AUC	0,989 ± 0,006	0,967 ± 0,037	0,920 ± 0,067
	Jl	0,946 ± 0,027	0,900 ± 0,075	0,721 ± 0,086
	DSC	0.972 ± 0.015	0.946 ± 0.044	0,838 ± 0,158
	HD (mm)	0.164 ± 0.174	0.373 ± 0.309	0,357 ± 0,277
Fold 3	AUC	0,980 ± 0,010	0,960 ± 0,035	0,913 ± 0,061
	Jl	0,940 ± 0,026	0,884 ± 0,066	0,701 ± 0,113
	DSC	0,969 ± 0,014	0.967 ± 0.039	0,824 ± 0,203
	HD (mm)	0,182 ± 0,127	0.393 ± 0.266	0,425 ± 0,387
Fold 4	AUC	0,983 ± 0,016	0,953 ± 0,037	0,887 ± 0,064
	Jl	0,924 ± 0,040	0,854 ± 0,080	0,602 ± 0,238
	DSC	0,960 ± 0,023	0,919 ± 0,050	0,748 ± 0,389
	HD (mm)	0,252 ± 0,232	0,526 ± 0,310	0,484 ± 0,298
Fold 5	AUC	0,987 ± 0,009	0,951 ± 0,051	0,884 ± 0,094
	Jl	0,941 ± 0,031	0,882 ± 0,098	0,677 ± 0,485
	DSC	0,969 ± 0,018	0,934 ± 0,061	0,773 ± 0,256
	HD (mm)	0,190 ± 0,214	0,413 ± 0,374	0,516 ± 0,788
Media	AUC	0,984 ± 0,004	0,956 ± 0,008	0,898 ± 0,018
	Jl	0,937 ± 0,008	0,880 ± 0,017	0,722 ± 0,040
	DSC	0,969 ± 0,006	0,952 ± 0,029	0,797 ± 0,037
	HD (mm)	0.242 ± 0.107	0.455 ± 0.088	0.431 ± 0.069

Tabla 4.5: Métricas de las tres etiquetas en imágenes con *stent*.

(en negrita). En caso de los *fold* 1 y 4 se tienen unos resultados peores en cuanto a todas las métricas, mientras que los mejores resultados son los del *fold* 2 para el lumen y el *fold* 3 para la pared vascular, teniendo este último resultados similares para la Distancia Hausdorff al *fold* 2, pero considerablemente mejores para el Índice de Dice.

La Figura 4.5 representa los diagramas de violín para las distintas etiquetas con las correspondientes métricas en el *fold* 2. En la Figura 4.5a se tiene el AUC, y en este caso es la métrica que mejor representa las prestaciones del modelo a la hora de segmentar imágenes con *stent*. Esto es, se obtienen muy buenas segmentaciones del lumen, pero las correspondientes a la pared vascular no son tan buenas. Por otro lado, la ROI obtiene resultados mucho más dispersos que en el caso de las imágenes con *stent*, debido a que es mucho más sensible a los errores de las etiquetas.

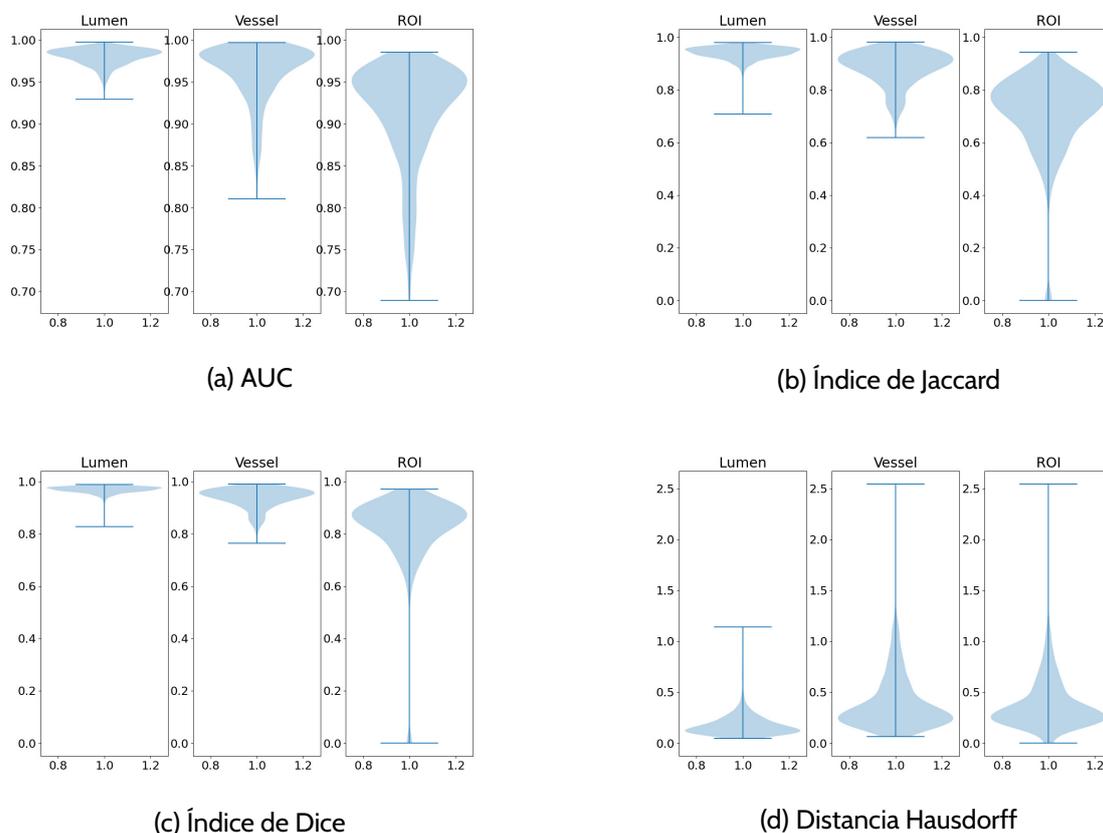


Figura 4.5: Diagramas de violín para el *fold* 2 para las distintas métricas en imágenes con *stent*.

Por último, en la Figura 4.6 encontramos los resultados gráficos correspondientes a la predicción y la etiqueta sobre el *frame* en cuestión. Las etiquetas del lumen y la pared vascular se encuentran delineadas en colores lima y rojo, respectivamente, mientras que las predicciones correspondientes al lumen y la pared vascular se encuentran en cian y amarillo, respectivamente.

En la primera fila se tienen algunas de las peores segmentaciones del modelo, mientras que en la segunda se tienen algunas de las mejor consideradas. Es de destacar que, en caso de estas imágenes y a diferencia de las de la Figura 4.4, la pared vascular generalmente no se ve de forma tan clara, provocando una segmentación imprecisa del algoritmo. En las Figuras 4.6d, 4.6e y 4.6f se aprecian de una mejor manera los límites de la pared vascular, correspondiéndose con mejores segmentaciones de la pared vascular, mientras que se puede observar todo lo contrario en las Figuras 4.6a, 4.6b y 4.6c.

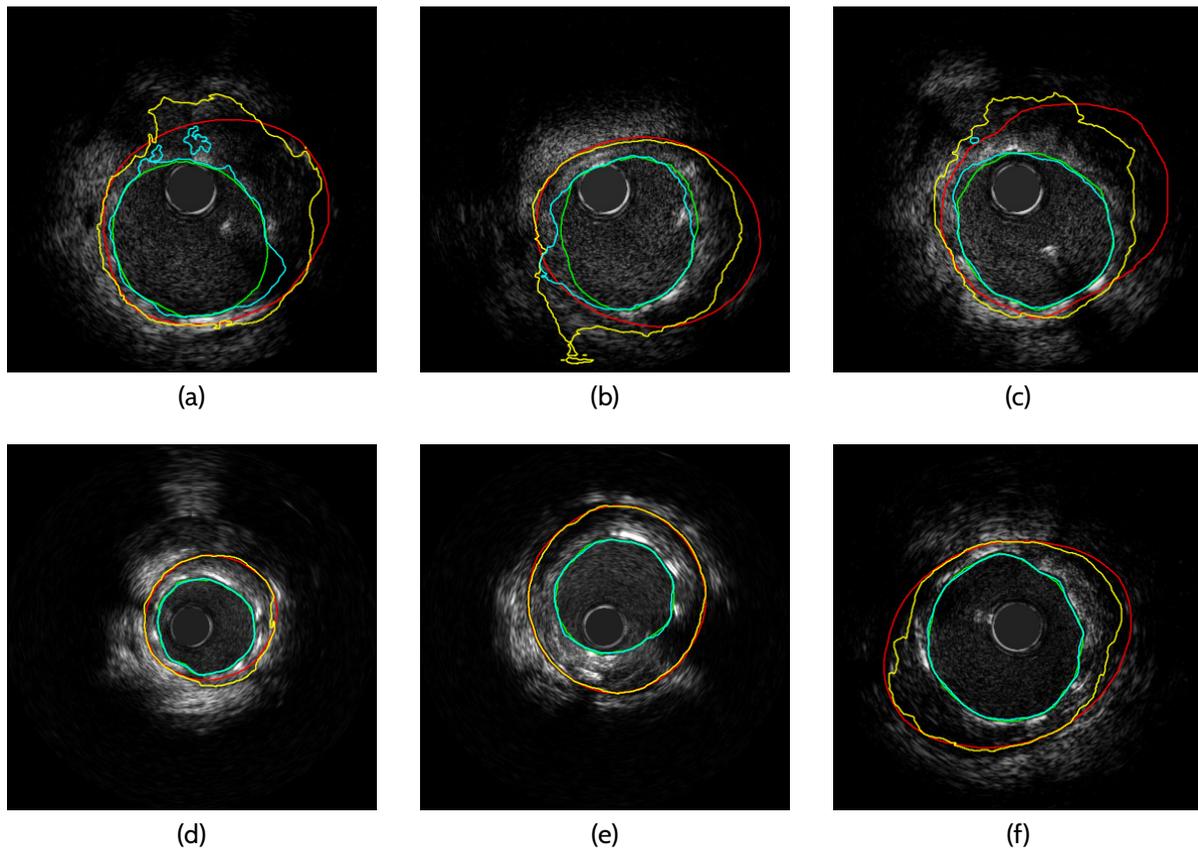


Figura 4.6: Contornos de predicciones con *stent* del modelo final optimizado.

4.5. Comparativa con otros métodos

Por último, para tener una idea de la relevancia de este trabajo, se realiza una comparativa con las métricas seleccionadas de los distintos trabajos que se consultaron, varios de ellos siendo considerados del *state-of-the-art*.

Hay que tener en cuenta que esta comparativa solo debe servir de método orientativo, ya que los datos con los que se han entrenado los distintos modelos son distintos, y varios de ellos incluso han sido adquiridos a distintas frecuencias y resoluciones. La comparativa se encuentra en la Tabla 4.6 y en el caso del método propuesto, que se encuentra en negrita, la

desviación estándar se corresponde con la media de la desviación estándar entre imágenes de los 5 *folds*, a diferencia de apartados anteriores.

Método	Métrica	Lumen	Pared vascular
Análisis BLOB [17]	Jl	0,9570	0,9670
	HD	0,5259	0,6304
<i>Supapixel Clustering</i> [18]	Jl	$0,78 \pm 0,10$	$0,83 \pm 0,10$
	HD (mm)	$1,18 \pm 0,70$	$1,20 \pm 0,66$
<i>MFCNN for lumen</i> [20]	Jl	0,925, [0,911, 0,948]	–
<i>8 layer U-Net</i> [22]	Jl	0,937	0,804
IVUS-Net [24]	Jl	$0,90 \pm 0,06$	$0,86 \pm 0,11$
	HD (mm)	$0,26 \pm 0,25$	$0,48 \pm 0,44$
<i>LV segmentation</i> [21]	Jl	0,913, [0,882, 0,935]	0,940, [0,917, 0,957]
	HD (mm)	0,196, [0,146, 0,275]	0,163, [0,122, 0,234]
<i>CNN for LV segmentation</i> [25]	Jl	0,911	0,933
	HD (mm)	0,336	0,367
Método propuesto	Jl	$0,909 \pm 0,054$	$0,925 \pm 0,058$
	HD (mm)	$0,207 \pm 0,197$	$0,278 \pm 0,240$

Tabla 4.6: Comparativa entre distintos métodos.

Los dos primeros métodos se corresponden con técnicas de *Machine Learning*, siendo por tanto métodos menos robustos. Respecto al análisis BLOB [17], las métricas proporcionadas por los autores son considerablemente elevadas. Sin embargo, hay que tener en cuenta que la resolución de las imágenes utilizadas por la referencia es menor que la utilizada en este trabajo. Esto provoca que el Índice de Jaccard sea más elevado al no haber detalles en los píxeles. Además, al presentar la Distancia Hausdorff de manera adimensional y depender esta de la resolución de la imagen, comparación es incierta. En cuanto al segundo trabajo [18], se tienen unas métricas inferiores a las del trabajo que aquí se presenta, siendo también mucho más pesado computacionalmente.

El resto de trabajos se centran en técnicas de *Deep Learning*, y tienen mejores métricas que la anterior referencia [18], debido a que los trabajos de *Deep Learning* consiguen extraer de mejor manera los patrones dentro de los datos [19]. En primer lugar, hay un trabajo que se centra exclusivamente en la segmentación del lumen. Para ello utilizan un total de 20586 *frames*, habiendo aplicado *gating* para la adquisición de las mismas. El *gating* es una técnica que utiliza el electrocardiograma del paciente para adquirir la imagen de IVUS, de manera que se elimine el ruido debido a las contracciones y expansiones de la pared vascular por el latido del corazón. Posteriormente, aplican un modelo de *Deep Learning* basado en U-Net [47] para segmentar los *frames*. El resultado es una segmentación bastante precisa del lumen en términos de índice de Jaccard, teniendo en cuenta que el modelo es capaz de segmentar exclusivamente el lumen.

En otro trabajo [22] se presentó una modificación de U-Net [47], de manera que se

extendiera a 8 capas ocultas en lugar de 6, para segmentación de imágenes de IVUS. En este caso, el total de *frames* ascendía a 591. Esto provocaba que el conjunto de test no fuera lo suficientemente grande, y cause una diferencia tan notable en el Índice de Jaccard del lumen y la pared vascular.

Varios autores decidieron tomar referencia de las redes de segmentación con mejores prestaciones en distintas aplicaciones para crear IVUS-Net [24]. En este caso, los datos utilizados por los autores también son bastante escasos, formando un total de 469 *frames*. Los resultados presentados en el trabajo son notablemente más equilibrados que los anteriores, y podría ser a que los conjuntos de entrenamiento y test fueron específicamente diseñados por los cardiólogos y, por tanto, estén mejor balanceados. De cualquier manera, esta red junto con la U-Net [47] y Deeplabv3+ [26] fueron exploradas en un único trabajo bajo las mismas condiciones y datos [25]. En este caso, se utilizó un total de 6516 *frames* para evaluar a los distintos modelos, encontrando a Deeplabv3+ [26] con ResNet101 [48] como el modelo que mejores prestaciones proporcionaba.

Por último, se presentó un trabajo que pretendía segmentar lumen y pared vascular utilizando U-Net [47] y aplicando una novedosa técnica de aumento de datos [21]. Para ello, era necesario utilizar *gating* en la adquisición de los datos, ya que se utilizaba la información de los *frames* sin etiquetar para interpolar las segmentaciones de los cardiólogos. El *gating* hace que haya una menor diferencia entre *frames* adyacentes, provocando así una interpolación más precisa que de no haberlos adquirido con *gating*. Tras segmentar el modelo utilizando U-Net, se utiliza un GPR para regularizar el contorno de la predicción, consiguiendo unos resultados satisfactorios. El total de *frames* utilizado en este trabajo es de 13435, y no se especifica el número total de *frames* utilizado tras el aumento de datos. La resolución utilizada también difiere ligeramente de la que se presenta en el actual trabajo, siendo de 500×500 . Sin embargo, al presentar la Distancia Hausdorff en milímetros, esta diferencia debería ser mínima en el Índice de Jaccard.

5 Discusión

En este manuscrito, se presenta un trabajo para la segmentación del lumen y la pared vascular en imágenes de IVUS compuesto por un modelo de aprendizaje profundo, Deeplabv3+ con *Modified Aligned Xception* como columna vertebral del modelo. Esta configuración es la que mejores resultados dio en el artículo original de Deeplabv3+ [26], resultando en una red de segmentación del *state-of-the-art* en cuanto a imágenes naturales se refiere.

Las predicciones del modelo son postprocesadas mediante apertura y cierre morfológicos con distinto tamaño de elemento estructurante, en función del tamaño de la etiqueta. También se aplicó una búsqueda de hiperparámetros mediante optimización bayesiana con resultados satisfactorios, y se incluyó la deformación con coordenadas polares como método de aumento de datos.

En primer lugar, con respecto a la segmentación de imágenes sin *stent*, se obtuvo un Índice de Dice de 0.952 para el lumen y 0.960 para la pared vascular. Por otro lado, se consiguió una Distancia Hausdorff media de 0.207 mm y 0.278 mm para el lumen y la pared vascular, respectivamente. Para imágenes con *stent*, se obtuvo un Índice de Dice de 0.967 para el lumen y 0.940 para la pared vascular. Las Distancias Hausdorff en este caso se encuentran en 0.242 mm y 0.455 mm para el lumen y la pared vascular, respectivamente.

Aunque los resultados obtenidos son satisfactorios, hay ciertos casos erráticos donde el modelo no tuvo un resultado correcto, como se muestra en la Figura 4.4. El algoritmo, en varios casos, predice el lumen y la pared vascular prácticamente superpuestos, como muestran las Figuras 4.4a y 4.4c. En la Figura 4.4b, que se corresponde con una bifurcación, el algoritmo tiene un resultado errático, mostrando unos contornos que difieren en mayor medida de los de la etiqueta en este caso. Las imágenes de este tipo son las más complicadas de segmentar, principalmente debido a que el lumen tiene una forma irregular y, generalmente, elíptica en lugar de circular.

Como indicaban otros trabajos, en la sección 4.1, se intentó entrenar el modelo con imágenes de distinta dimensionalidad. En primer lugar, se entrenó el modelo con imágenes

multiespectrales en 3D, compuestas por el *frame* etiquetado por el cardiólogo y dos más, seleccionados de entre un vecindario definido bajo distintos criterios. Se obtuvo un resultado distinto al marcado en [25], mostrando que usar exclusivamente el mismo *frame* que la etiqueta es la mejor forma de proceder bajo las condiciones del experimento presentado en este trabajo.

En primer lugar, esta diferencia hallada podría deberse a los datos utilizados en este trabajo. Durante la adquisición de estos no se aplicó *gating*, es decir, la adquisición no estuvo controlada mediante electrocardiograma, a diferencia de los utilizados en la referencia [25]. Esta técnica se utiliza para intentar minimizar las diferencias entre *frames* adyacentes debido a las contracciones y expansiones del lumen y la pared vascular por el latido del corazón.

Por otra parte, esta diferencia podría deberse al hecho de haber realizado experimentos con una columna vertebral distinta. Recordemos que los experimentos presentados en este trabajo están realizados con *Modified Alligned Xception*, mientras que la referencia utiliza *ResNet 101* para ello. Sin embargo, estas diferencias no pueden concluirse como las causantes del resultado distinto, y más investigación es requerida en este sentido.

En la sección 4.2, se incluyó la deformación polar como técnica de aumento de datos. Se vio que esta consigue un entrenamiento más robusto que la deformación elástica, modificando por tanto las imágenes aumentadas de una manera más natural. Esto tiene un impacto directo en las etapas de validación y testeo, mejorando las prestaciones finales del modelo sobre estos conjuntos.

En la sección 4.3, se añadió al modelo un post-procesado para regularizar los contornos e intentar resolver ciertas predicciones que contenían segmentaciones no cerradas en imágenes sin *stent*. El resultado fue una mejora notable en todas las métricas, especialmente la Distancia Hausdorff de la pared vascular. Por otro lado, este mismo flujo de trabajo se aplicó a imágenes con *stent* en la sección 4.4. Se esperaba obtener unos resultados con prestaciones inferiores al caso de tener imágenes sin *stent* debido a que el número de estas últimas era menor. Además, las imágenes con *stent* contienen generalmente un lumen más circular y regular debido a la inclusión de la malla metálica. Sin embargo, esto provoca que las ondas emitidas por la sonda no penetren tan profundamente en el tejido, haciendo menos visible la pared vascular y, por tanto, dificultando su segmentación. Esto provoca que el algoritmo no segmente adecuadamente la pared vascular en aquellos casos en los que la *Intima* es más grande al haber mayor distancia entre lumen y pared vascular.

Es de notar que en la Figuras 4.3 y 4.5 se observan los diagramas de violín conteniendo las distintas métricas del modelo para *frames* con y sin *stent*. Todas las etiquetas de la ROI muestran un valor más disperso, debido a que el área de la imagen que se corresponde con la ROI es considerablemente más sensible a los errores en los contornos del lumen

y la pared vascular. Por esta razón, se considera una muy buena forma de condensar las prestaciones de un modelo en un único valor a la hora de compararlo con otros modelos similares, teniendo en cuenta que estos serán más restrictivos.

A la hora de comparar el modelo con trabajos que abordan el mismo problema, lo que se corresponde con la sección 4.5, se encuentra que los resultados obtenidos en este trabajo son similares a los vistos en la literatura, incluso siendo mejores en ciertos casos. Principalmente, el modelo descrito en este trabajo destaca por la robustez de sus contornos a pesar de tener un post-procesado relativamente simple, llegando a tener prestaciones similares a otros trabajos en los que se utilizan técnicas más avanzadas, como Regresores de Procesos Gaussianos [20] [21]. En otros trabajos se ha optado por aplicar técnicas novedosas más complicadas en cuanto a aumento de datos [25].

En el caso de este trabajo, los datos utilizados han sido segmentados por distintos cardiólogos, lo que hace que haya un cierto ruido debido a la varianza inter-expertos. También se ha visto que aquellos trabajos en los que se ha utilizado el *gating* en la adquisición de los datos han obtenido mejores resultados, seguramente debido a la mayor estabilidad de los *frames* en el *pullback* adquirido. El hecho de utilizar catéteres de distintas empresas también añade pequeñas modificaciones en el brillo y contraste, que no deberían afectar a este tipo de modelos pero sí a otros que no basen su funcionamiento en la convolución.

Como resultado de este Trabajo de Fin de Máster, se ha obtenido un modelo particularmente robusto, capaz de segmentar automáticamente imágenes con un tiempo de procesado considerablemente corto, y que podría ser de gran ayuda para los cardiólogos en la toma de decisiones durante PCI en función de los contornos del lumen y la pared vascular.

5.1. Futuras líneas

La naturaleza de adquisición de los *pullbacks* permite tener todas las imágenes en una pila, de manera que se puedan utilizar previas muestras en la predicción de siguientes. En este sentido, las Memorias a líneas temporales a corto plazo de larga duración (Long Short-Term Memory, LSTM), se pueden utilizar para predecir muestras siguientes teniendo en cuenta predicciones anteriores, lo que tendría un resultado bastante positivo debido a la continuidad de estas imágenes.

Por otro lado, al tener todo el corte longitudinal del *pullback*, se puede utilizar este para intentar localizar zonas problemáticas que requieran una observación detallada por parte del cardiólogo, o ayudar de algún modo a la segmentación aportando información adicional de continuidad del vecindario.

La identificación de trombos es compleja en imágenes de IVUS debido a que estos no destacan por las características de adquisición de la imagen. Por ello, la detección asistida de estos cuerpos facilitaría la identificación de trombosis durante PCI, no existiendo hoy en día un sistema capaz de identificarlo en imágenes de IVUS.

El tipo de sistemas de ayuda a los cardiólogos presentado en este trabajo podría aplicarse a otras modalidades de imagen. Como se ha visto, actualmente hay combinaciones de modalidades de imagen que pueden proporcionar un mejor detalle a las imágenes, como IVUS + OCT, que permite mantener el buen detalle del lumen mientras se observa mejor el límite con la pared virtual. Otra alternativa es IVUS + NIRS, que permite obtener la probabilidad de ver lípidos en estas imágenes.

El número de bases de datos disponible es bastante escaso, siendo la mayoría de ellas de uso privado. Con este sentido, diversos son los trabajos que presentan *Federated Learning* para solventar el problema de compartir datos entre distintos hospitales o centros médicos. Este método consiste en un cambio de paradigma y, en lugar de compartir los datos para utilizar distintos modelos sobre ellos, se utilizarían distintos datos para el mismo modelo, de manera que se tenga una mejor referencia de las prestaciones que dicho modelo puede alcanzar. En su defecto, la creación de nuevas y más extensas bases de datos de imágenes de IVUS para segmentación de lumen y pared vascular también podría ser una solución.

Bibliografía

- [1] C. Clinic, “Coronary artery disease,” 2022.
- [2] M. A. Khan, M. J. Hashim, H. Mustafa, M. Y. Baniyas, S. K. B. M. A. Suwaidi, R. AlKatheeri, F. M. K. Alblooshi, M. E. A. H. Almatrooshi, M. E. H. Alzaabi, R. S. A. Darmaki, and S. N. A. H. Lootah, “Global epidemiology of ischemic heart disease: Results from the global burden of disease study,” *Int J Cardiovasc Imaging*, vol. 27, no. 2, pp. 215–224, 2011.
- [3] M. Plus, “Enfermedad de las arterias coronarias,” 2022.
- [4] F. T. W. Groenland, K. D. Mahmoud, T. Neleman, A. C. Z. des Plantes, A. Scoccia, J. Ligthart, K. T. Witberg, R.-J. Nuis, W. K. den Dekker, J. M. Wilschut, R. Diletti, F. Zijlstra, I. Kardys, P. Cummins, N. M. V. Mieghem, and J. Daemen, “Tissue characterisation and primary percutaneous coronary intervention guidance using intravascular ultrasound: rationale and design of the spectrum study,” *Open Heart*, vol. 9, no. 1, 2022.
- [5] U. Jin, H.-S. Lim, and S.-J. Tahk, “Longitudinal stent-strut injury at the distal end of a newer-generation drug-eluting stent,” *Korean Circulation Journal*, vol. 48, no. 2, 2017.
- [6] P. Gong, M. Almasian, G. van Soest, D. de Bruin, T. van Leeuwen, D. Sampson, and D. Faber, “Parametric imaging of attenuation by optical coherence tomography: review of models, methods, and clinical translation,” *Journal of Biomedical Optics*, vol. 25, no. 4, 2020.
- [7] H. M. García-García, B. D. Gogas, P. W. Serruys, and N. Bruining, “Ivus-based imaging modalities for tissue characterization: similarities and differences,” *Cureus*, vol. 12, no. 7, 2020.
- [8] T. Roleder, J. Jąkała, G. L. Kałuża, Łukasz Partyka, K. Proniewska, E. Pociask, W. Zasada, W. Wojakowski, Z. Gašior, and D. Dudek, “The basics of intravascular optical coherence tomography,” *Postepy Kardiol Interwencyjnej*, vol. 11, no. 2, 2015.
- [9] Z. A. Ali, K. K. Galougahi, G. S. Mintz, A. Maehara, R. A. Shlofmitz, and A. Mattesini, “Intracoronary optical coherence tomography: state of the art and future directions,” *EuroIntervention*, vol. 17, no. 2, 2021.

- [10] B. H. Li, A. S. Leung, A. Soong, C. E. Munding, H. Lee, A. S. Thind, N. R. Munce, G. A. Wright, C. H. Rowsell, V. X. Yang, B. H. Strauss, F. S. Foster, and B. K. Courtney, "Hybrid intravascular ultrasound and optical coherence tomography catheter for imaging of coronary atherosclerosis," *Catheterization and Cardiovascular Interventions*, vol. 81, no. 3, 2012.
- [11] M. Ono, H. Kawashima, H. Hara, C. Gao, R. Wang, N. Kogame, K. Takahashi, P. Chichareon, R. Modolo, M. Tomaniak, J. J. W. J. J. Piek, I. Mori, B. K. Courtney, W. Wijns, F. Sharif, C. Bourantas, Y. Onuma, and P. W. Serruys, "Advances in ivus/oct and future clinical perspective of novel hybrid catheter system in coronary imaging," *Frontiers in Cardiovascular Medicine*, vol. 7, 2020.
- [12] K. O. Kuku, M. Singh, Y. Ozaki, K. Dan, C. Chezar-Azerrad, R. Waksman, and H. M. Garcia-Garcia, "Near-infrared spectroscopy intravascular ultrasound imaging: State of the art," *Frontiers in Cardiovascular Medicine*, vol. 7, no. 107, 2020.
- [13] R. Waksman, C. D. Mario, R. Torguson, Z. A. Ali, V. Singh, W. H. Skinner, A. K. Artis, T. T. Cate, E. Powers, C. Kim, E. Regar, S. C. Wong, S. Lewis, J. Wykrzykowska, S. Dubbe, S. Kazziha, M. van der Ent, P. Shah, P. E. Craig, Q. Zou, P. Kolm, H. B. Brewer, and H. M. Garcia-Garcia, "Identification of patients and plaques vulnerable to future coronary events with near-infrared spectroscopy intravascular ultrasound imaging: a prospective, cohort study," *Lancet*, vol. 394, 2019.
- [14] N. C. Giglio, S. Farrow, K. Sreeram, and S. Grainger, "Intravascular near-infrared spectroscopy predicts heart attacks," 2023.
- [15] K. Terada, T. Kubo, T. Kameyama, Y. Matsuo, Y. Ino, H. Emori, D. Higashioka, Y. Katayama, A. K. Khalifa, M. Takahata, K. Shimamura, Y. Shiono, A. Tanaka, T. Hozumi, R. D. Madder, and T. Akasaka, "Nirs-ivus for differentiating coronary plaque rupture, erosion, and calcified nodule in acute myocardial infarction," *JACC: Cardiovascular Imaging*, vol. 7, 2021.
- [16] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. University of Tennessee, TN, USA: Pearson, 3rd ed., 2007.
- [17] *Automatical IVUS Image Segmentation and Detection Using BLOB Analysis*, vol. 1 of *Lecture Notes in Networks and Systems*, Springer, 2022.
- [18] M. Xia, W. Yan, Y. Huang, Y. Guo, G. Zhou, and Y. Wang, "Ivus image segmentation using superpixel-wise fuzzy clustering and level set evolution," *Applied Sciences*, vol. 9, no. 22, 2019.
- [19] X. Liu, L. Song, S. Liu, and Y. Zhang, "A review of deep-learning-based medical image segmentation methods," *Sustainability*, vol. 13, no. 3, 2021.

- [20] P. G. P. Ziemer, C. A. Bulant, J. I. Orlando, G. D. M. Talou, L. A. M. Álvarez, C. G. Bezerra, P. A. Lemos, H. M. García-García, and P. J. Blanco, "Automated lumen segmentation using multi-frame convolutional neural networks in intravascular ultrasound datasets," *European Heart Journal - Digital Health*, vol. 1, no. 1, 2020.
- [21] P. J. Blanco, P. G. Ziemer, C. A. Bulant, Y. Ueki, R. Bass, L. Räber, P. A. Lemos, and H. M. García-García, "Fully automated lumen and vessel contour segmentation in intravascular ultrasound datasets," *Medical Image Analysis*, vol. 75, 2022.
- [22] L. Dong, W. Jiang, W. Lu, J. Jiang, Y. Zhao, X. Song, X. Leng, H. Zhao, J. Wang, C. Li, and J. Xiang, "Automatic segmentation of coronary lumen and external elastic membrane in intravascular ultrasound images using 8-layer u-net," *BioMedical Engineering OnLine*, vol. 20, no. 16, 2021.
- [23] L. Bargsten, S. Raschka, and A. Schlaefer, "Capsule networks for segmentation of small intravascular ultrasound image datasets," *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, 2021.
- [24] J. Yang, L. Tong, M. Faraji, and A. Basu, "Ivus-net: An intravascular ultrasound segmentation network," *Proc. 18th International Conference on Soft Materials (ICSM, Busan, Korea, December 2018)*, vol. 1, 2018.
- [25] H. Du, L. Ling, W. Yu, P. Wu, Y. Yang, M. Chu, J. Yang, W. Yang, and S. Tu, "Convolutional networks for the segmentation of intravascular ultrasound images: Evaluation on a multicenter dataset," *Computer Methods and Programs in Biomedicine*, vol. 215, 2022.
- [26] *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation*, vol. 11211 of *European Conference on Computer Vision*, Springer, 2018.
- [27] S.-H. Tsang, "Review: Xception — with depthwise separable convolution, better than inception-v3 (image classification)," 2018.
- [28] *Adam: A Method for Stochastic Optimizatio*, ICLR 2015, 2014.
- [29] C. Collins, "Parameter hyperparameter tuning with bayesian optimization," 2019.
- [30] J. Brownlee, "Scikit-optimize for hyperparameter tuning in machine learning," 2020.
- [31] A. Mohan, "Bayesian optimization and hyperparameter tuning," 2021.
- [32] W. Koehrsen, "A conceptual explanation of bayesian hyperparameter optimization for machine learning," 2018.
- [33] M. A. Gamboa, "Hyperparameter optimization using bayesian optimization," 2019.

- [34] R. Chandradevan, “Shallow understanding on bayesian optimization,” 2017.
- [35] S. Boutadghart, “Bayesian optimization,” 2017.
- [36] J. Torres, *Introduction to Deep Learning*. Gran Via de les Corts Catalanes, Barcelona: Marcombo, 1st ed., 2020.
- [37] *Healthy Vessel Wall Detection Using U-Net in Optical Coherence Tomography*, Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting, Springer-Verlag, 2019.
- [38] *Best practices for convolutional neural networks applied to visual document analysis*, IEEE, 2003.
- [39] T. Neleman, L. J. C. van Zandvoort, M. N. T. Forero, K. Masdjedi, J. M. R. Ligthart, K. T. Witberg, F. T. W. Groenland, P. Cummins, M. J. Lenzen, E. Boersma, R.-J. Nuis, W. K. den Dekker, R. Diletti, J. Wilschut, F. Zijlstra, N. M. V. Mieghem, and J. Daemen, “Ffr-guided pci optimization directed by high-definition ivus versus standard of care: The ffr react trial,” *JACC: Cardiovascular Interventions*, vol. 15, no. 16, 2022.
- [40] T. Developers, “Tensorflow,” Aug. 2023.
- [41] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.
- [42] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [43] “Curva roc,” 2023.
- [44] “Jaccard index,” 2023.
- [45] “Hausdorff distance,” 2023.
- [46] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burrowski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.

- [47] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention*, 2015.
- [48] *Deep Residual Learning for Image Recognition*, 2016.